PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



INTERNATIONAL APPLICATION PUBLISH	HED U	JNDER THE PATENT COOPERATION TREATY (PCT)
(51) International Patent Classification 7:		(11) International Publication Number: WO 00/39303
C12N 15/49, C07K 14/16	A2	(43) International Publication Date: 6 July 2000 (06.07.00)
(21) International Application Number: PCT/US (22) International Filing Date: 30 December 1999 ((30) Priority Data:	30.12.9 8) U 9) U	BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZA, ZW ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAP patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR
Street, Emeryville, CA 94608 (US). (72) Inventors: BARNETT, Susan; Chiron Corporation, 4 ton Street - R440, Emeryville, CA 94608 (US). H Karin; Chiron Corporation, 4560 Horton Street Emeryville, CA 94608 (US). MARTIN, Eric; Ch poration, 4560 Horton Street - R440, Emeryville, C (US).	ARTO - R44 iron Co	G, Published O, Without international search report and to be republished upon receipt of that report.
(74) Agents: DOLLARD, Anne, S.; Chiron Corporation, tual Property - R440, P.O. Box 8097, Emery 94662-8097 (US) et al.	Intelle ville, C	C- A
(54) Title: MODIFIED HIV ENV POLYPEPTIDES		
(57) Abstract		
		disclosed. The Env polypeptides are modified so as to expose at least par prevention using the polynucleotides and polypeptides are also provided

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
ΑT	Austria	FR	Prance	LU	Luxembourg	SN	Senegal
ΑU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
ΑZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin .	1E	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ.	Uzbekistan
CF	Central African Republic	JР	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	ΥU	Yugoslavia
СН	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

MODIFIED HIV ENV POLYPEPTIDES

Technical Field

5

10

15

20

25

30

The invention relates generally to modified HIV envelope (Env) polypeptides which are useful as immunizing agents or for generating an immune response in a subject, for example a cellular immune response or a protective immune response. More particularly, the invention relates Env polypeptides such as gp120, gp140 or gp160, wherein at least one of the native β -sheet configurations has been modified. The invention also pertains to methods of using these polypeptides to elicit an immune response against a broad range of HIV subtypes.

Background of the Invention

The human immunodeficiency virus (HIV-1, also referred to as HTLV-III, LAV or HTLV-III/LAV) is the etiological agent of the acquired immune deficiency syndrome (AIDS) and related disorders. (see, e.g., Barre-Sinoussi, et al., (1983) Science 220:868-871; Gallo et al. (1984) Science 224:500-503; Levy et al., (1984) Science 225:840-842; Siegal et al., (1981) N. Engl. J. Med. 305:1439-1444). AIDS patients usually have a long asymptomatic period followed by the progressive degeneration of the immune system and the central nervous system. Replication of the virus is highly regulated, and both latent and lytic infection of the CD4 positive helper subset of T-lymphocytes occur in tissue culture (Zagury et al., (1986) Science 231:850-853). Molecular studies of HIV-1 show that it encodes a number of genes (Ratner et al., (1985) Nature 313:277-284; Sanchez-Pescador et al., (1985) Science 227:484-492), including three structural genes -- gag, pol and env -- that are common to all retroviruses. Nucleotide sequences from viral genomes of other retroviruses, particularly HIV-2 and simian immunodeficiency viruses, SIV (previously referred to as STLV-III), also contain these structural genes. (Guyader et al., (1987) Nature 326:662-669; Chakrabarti et al., (1987) Nature

The envelope protein of HIV-1, HIV-2 and SIV is a glycoprotein of about 160 kd (gp160). During virus infection of the host cell, gp160 is cleaved by host cell proteases to form gp120 and the integral membrane protein, gp41. The gp41 portion is anchored in the

5

10

15

20

25

30

membrane bilayer of virion, while the gp120 segment protrudes into the surrounding environment. gp120 and gp41 are more covalently associated and free gp120 can be released from the surface of virions and infected cells.

As depicted in Figure 1, crystallography studies of the gp120 core polypeptide indicate that this polypeptide is folded into two major domains having certain emanating structures. The inner domain (inner with respect to the N and C terminus) features a two-helix, two-stranded bundle with a small five-stranded β -sandwich at its termini-proximal end and a projection at the distal end from which the V1/V2 stem emanates. The outer domain is a staked double barrel that lies along side the inner domain so that the outer barrel and inner bundle axes are approximately parallel. Between the distal inner domain and the distal outer domain is a four-stranded bridging sheet which holds a peculiar minidomain in contact with, but distinct from, the inner, the outer domain, and the V1/V2 domain. The bridging sheet is composed of four β -strand structures (β -3, β -2, β -21, β -20, shown in Figure 1). The bridging region can be seen in Figure 1 packing primarily over the inner domain, although some surface residues of the outer domain, such as Phe 382, reach into the bridging sheet to form part of its hydrophobic core.

The basic unit of the β -sheet conformation of the bridging sheet region is the β -strand which exists as a less tightly coiled helix, with 2.0 residues per turn. The β -strand conformation is only stable when incorporated into a β -sheet, where hydrogen bonds with close to optimal geometry are formed between the peptide groups on adjacent β -strands; the dipole moments of the strands are also aligned favorably. Side chains from adjacent residues of the same strand protrude from opposite sides of the sheet and do not interact with each other, but have significant interactions with their backbone and with the side chains of neighboring strands. For a general description of β -sheets, see, e.g., T.E. Creighton, <u>Proteins: Structures and Molecular Properties</u> (W.H. Freeman and Company, 1993); and A.L. Lehninger, <u>Biochemistry</u> (Worth Publishers, Inc., 1975).

The gp120 polypeptide is instrumental in mediating entry into the host cell. Recent studies have indicated that binding of CD4 to gp120 induces a conformational change in Env that allows for binding to a co-receptor (e.g., a chemokine receptor) and subsequent entry of the virus into the cell. (Wyatt, R., et al. (1998) *Nature* 393:705-711; Kwong, P., et al. (1998) *Nature* 393:648-659). Referring again to Figure 1, CD4 is bound into a depression formed at the interface of the outer domain, the inner domain and the bridging sheet of gp120.

Immunogenicity of the gp120 polypeptide has also been studied. For example, individuals infected by HIV-1 usually develop antibodies that can neutralize the virus in *in vitro* assays, and this response is directed primarily against linear neutralizing determinants in the third variable loop of gp120 glycoprotein (Javaherian, K., et al. (1989) *Proc. Natl. Acad. Sci.* 86:6786-6772; Matsushita, M., et al. (1988) *J. Virol.* 62:2107-2144; Putney, S., et al. (1986) *Science* 234:1392-1395; Rushe, J. R., et al. (1988) *Proc. Nat. Acad. Sci. USA* 85: 3198-3202.). However, these antibodies generally exhibit the ability to neutralize only a limited number of HIV-1 strains (Matthews, T. (1986) *Proc. Natl. Acad. Sci. USA*. 83:9709-9713; Nara, P. L., et al. (1988) *J. Virol.* 62:2622-2628; Palker, T. J., et al. (1988) *Proc. Natl. Acad. Sci. USA*. 85:1932-1936). Later in the course of HIV infection in humans, antibodies capable of neutralizing a wider range of HIV-1 isolates appear (Barre-Sinoussi, F., et al. (1983) *Science* 220:868-871; Robert-Guroff, M., et al. (1985) *Nature* (London) 316:72-74; Weis, R., et al. (1985) *Nature* (London) 316:69-72; Weis, R., et al. (1986) *Nature* (London) 324:572-575).

Recent work done by Stamatatos et al (1998) AIDS Res Hum Retroviruses

14(13):1129-39, shows that a deletion of the variable region 2 from a HIV-1_{SF162} virus, which utilizes the CCR-5 co-receptor for virus entry, rendered the virus highly susceptible to serum-mediated neutralization. This V2 deleted virus was also neutralized by sera obtained from patients infected not only with clade B HIV-1 isolates but also with clade A, C, D and F HIV-1 isolates. However, deletion of the variable region 1 had no effect. Deletion of the variable regions 1 and 2 from a LAI isolate HIV-I_{IIIB} also increased the susceptibility to neutralization by monoclonal antibodies whose epitopes are located within the V3 loop, the CD4-binding site, and conserved gp120 regions (Wyatt, R., et al. (1995) J Virol. 69:5723-5733). Rabbit immunogenicity studies done with the HIV-1 virus with deletions in the V1/V2 and V3 region from the LAI strain, which uses the CXCR4 co-receptor for virus entry, showed no improvement in the ability of Env to raise neutralizing antibodies (Leu et al. (1998) AIDS Res. and Human Retroviruses. 14:151-155).

Further, a subset of the broadly reactive antibodies, found in most infected individuals, interferes with the binding of gp120 and CD4 (Kang, C.-Y., et al. (1991) *Proc. Natl. Acad. Sci. USA.* 88:6171-6175; McDougal, J. S., et al. (1986) *J. Immunol.* 137:2937-2944). Other antibodies are believed to bind to the chemokine receptor binding region after CD4 has bound to Env (Thali et al. (1993) *J. Virol.* 67:3978-3988). The fact that neutralizing

antibodies generated during the course of HIV infection do not provide permanent antiviral effect may in part be due to the generation of "neutralization escapes" virus mutants and to the general decline in the host immune system associated with pathogenesis. In contrast, the presence of pre-existing neutralizing antibodies upon initial HIV-1 exposure will likely have a protective effect.

5

10

15

20

25

30

It is widely thought that a successful vaccine should be able to induce a strong, broadly neutralizing antibody response against diverse HIV-1 strains (Montefiori and Evans (1999) AIDS Res. Hum. Ret. 15(8):689-698; Bolognesi, D.,P., et al. (1994) Ann. Int. Med. 8:603-611; Haynes, B., F., et al. (1996) Science; 271: 324-328.). Neutralizing antibodies, by attaching to the incoming virions, can reduce or even prevent their infectivity for target cells and prevent the cell-to-cell spread of virus in tissue culture (Hu et al. (1992) Science 255:456-459; Burton, D.,R. and Montefiori, D. (1997) AIDS 11(suppl. A): 587-598). However as described above, antibodies directed against gp120 do not generally exhibit broad antibody responses against different HIV strains.

Currently, the focus of vaccine development, from the perspective of humoral immunity, is on the neutralization of primary isolates that utilize the CCR5 chemokine coreceptor believed to be important in virus entry (Zhu, T., et al. (1993) *Science* 261:1179-1181; Fiore, J., et al. (1994) Virology; 204:297-303). These viruses are generally much more resistant to antibody neutralization than T-cell line adapted strains that use the CXCR4 coreceptor, although both can be neutralized *in vitro* by certain broadly and potent acting monoclonal antibodies, such as IgG1b12, 2G12 and 2F5 (Trkola, A., et al. (1995) *J. Virol*. 69:6609-6617; D'Sousa PM., et al (1997) *J. Infect. Dis.* 175:1062-1075). These monoclonal antibodies are directed to the CD4 binding site, a glycosylation site and to the gp41 fusion domain, respectively. The problem that remains, however, is that it is not known how to induce antibodies of the appropriate specificity by vaccination. Antibodies (Abs) elicited by gp120 glycoprotein from a given isolate are usually only able to neutralize closely related viruses generally from similar, usually from the same, HIV-1 subtype.

Despite the above approaches, there remains a need for Env antigens that can elicit an immunological response (e.g., neutralizing and/or protective antibodies) in a subject against multiple HIV strains and subtypes, for example when administered as a vaccine. The present invention solves these and other problems by providing modified Env polypeptides (e.g., gp120) to expose epitopes in or near the CD4 binding site.

Summary of the Invention

5

10

15

20

25

30

In accordance with the present invention, modified HIV Env polypeptides are provided. In particular, deletions and/or mutations are made in one or more of the 4-β antiparallel-bridging sheet in the HIV Env polypeptide. In this way, enough structure is left to allow correct folding of the polypeptide, for example of gp120, yet enough of the bridging sheet is removed to expose the CD4 groove, allowing an immune response to be generated against epitopes in or near the CD4 binding site of the Env polypeptide (e.g., gp120).

In one aspect, the invention includes a polynucleotide encoding a modified HIV Env polypeptide wherein the polypeptide has at least one modified (e.g., deleted or replaced) amino acid residue deleted in the region corresponding to residues 421 to 436 relative to HXB-2, for example the constructs depicted in Figures 6-29 (SEQ ID NOs:3 to 26). In certain embodiments, the polynucleotide also has the region corresponding to residues 124-198 of the polypeptide HXB-2 (e.g., V1/V2) deleted and at least one amino acid deleted or replaced in the regions corresponding to the residues 119 to 123 and 199 to 210, relative to HXB-2. In other embodiments, these polynucleotides encode Env polypeptides having at least one amino acid of the small loop of the bridging sheet (e.g., amino acid residues 427 to 429 relative to HXB-2) deleted or replaced. The amino acid sequences of the modified polypeptides encoded by the polynucleotides of the present invention can be based on any HIV variant, for example SF162.

In another aspect, the invention includes immunogenic modified HIV Env polypeptides having at least one modified (e.g., deleted or replaced) amino acid residue deleted in the region corresponding to residues 421 to 436 relative to HXB-2, for example a deletion or replacement of one amino acids in the small loop region (e.g., amino acid residues 427 to 429 relative to HXB-2). These polypeptides may have modifications (e.g., a deletion or a replacement) of at least one amino acid between about amino acid residue 420 and amino acid residue 436, relative to HXB-2 and, optionally, may have deletions or truncations of the V1 and/or V2 regions. The immunogenic, modified polypeptides of the present invention can be based on any HIV variant, for example SF162.

In another aspect, the invention includes a vaccine composition comprising any of the polynucleotides encoding modified Env polypeptides described above. Vaccine compositions comprising the modified Env polypeptides and, optionally, an adjuvant are also included in the invention.

In yet another aspect, the invention includes a method of inducing an immune response in subject comprising, administering one or more of the polynucleotides or constructs described above in an amount sufficient to induce an immune response in the subject. In certain embodiments, the method further comprises administering an adjuvant to the subject.

In another aspect, the invention includes a method of inducing an immune response in a subject comprising administering a composition comprising any of the modified Env polypeptides described above and an adjuvant. The composition is administered in an amount sufficient to induce an immune response in the subject.

In another aspect, the invention includes a method of inducing an immune response in a subject comprising

- (a) administering a first composition comprising any of the polynucleotides described above in a priming step and
- (b) administering a second composition comprising any of the modified Env polypeptides described above, as a booster, in an amount sufficient to induce an immune response in the subject. In certain embodiments, the first composition, the second composition or both the first and second compositions further comprise an adjuvant.

These and other embodiments of the subject invention will readily occur to those of skill in the art in light of the disclosure herein.

20

5

10

15

25

30

Brief Description of the Drawings

Figure 1 is a schematic depiction of the tertiary structure of the HIV-1_{HXB-2} Env gp120 polypeptide, as determined by crystallography studies.

Figures 2A-C depict alignment of the amino acid sequence of wild-type HIV-1_{HXB-2} Env gp160 polypeptide (SEQ ID NO:1) with amino acid sequence of HIV variants SF162 (shown as "162") (SEQ ID NO:2), SF2, CM236 and US4. Arrows indicate the regions that are deleted or replaced in the modified polypeptides. Black dots indicate conserved cysteine residues. The star indicates the position of the last amino acid in gp120.

Figures 3A-J depict alignment of nucleotide sequences of polynucleotides encoding modified Env polypeptides having V1/V2 deletions. The unmodified amino acid residues encoded by these sequences correspond to wildtype SF162 residues but are numbered relative to HXB-2.

Figures 4A-M depict alignment of nucleotide sequences of polynucleotides encoding modified Env polypeptides having deletions or replacements in the small loop. The unmodified amino acid residues encoded by these sequences correspond to wildtype SF162 residues but are numbered relative to HXB-2.

Figures 5A-N depict alignment of nucleotide sequences of polynucleotides encoding modified Env polypeptides having both V1/V2 deletions and, in addition, deletions or replacements in the small loop. The unmodified amino acid residues encoded by these sequences correspond to wildtype SF162 residues but are numbered relative to HXB-2.

5

10

15

25

30

Figure 6 depicts the nucleotide sequence of the construct designated Val120-Ala204 (SEQ ID NO:3).

Figure 7 depicts the nucleotide sequence of the construct designated Val120-Ile201 (SEQ ID NO:4).

Figure 8 depicts the nucleotide sequence of the construct designated Val120-Ile201B (SEQ ID NO:5).

Figure 9 depicts the nucleotide sequence of the construct designated Lys121-Val200 (SEQ ID NO:6).

Figure 10 depicts the nucleotide sequence of the construct designated Leu122-Ser199 (SEQ ID NO:7).

Figure 11 depicts the nucleotide sequence of the construct designated Val120-Thr202 (SEQ ID NO:8).

Figure 12 depicts the nucleotide sequence of the construct designated Trp427-Gly431 (SEQ ID NO:9).

Figure 13 depicts the nucleotide sequence of the construct designated Arg426-Gly431 (SEQ ID NO:10).

Figure 14 depicts the nucleotide sequence of the construct designated Arg426-Gly431B (SEQ ID NO:11).

Figure 15 depicts the nucleotide sequence of the construct designated Arg426-Lys432 (SEQ ID NO:12).

Figure 16 depicts the nucleotide sequence of the construct designated Asn425-Lys432 (SEQ ID NO:13).

Figure 17 depicts the nucleotide sequence of the construct designated Ile424-Ala433 (SEQ ID NO:14).

Figure 18 depicts the nucleotide sequence of the construct designated Ile423-Met434 (SEQ ID NO:15).

Figure 19 depicts the nucleotide sequence of the construct designated Gln422-Tyr435 (SEQ ID NO:16).

Figure 20 depicts the nucleotide sequence of the construct designated Gln422-Tyr435B (SEQ ID NO:17).

Figure 21 depicts the nucleotide sequence of the construct designated Leu122-Ser199;Arg426-Gly431 (SEQ ID NO:18).

Figure 22 depicts the nucleotide sequence of the construct designated Leu122-Ser199;Arg426-Lys432 (SEQ ID NO:19).

Figure 23 depicts the nucleotide sequence of the construct designated Leu122-Ser199; Trp427-Gly431 (SEQ ID NO:20).

Figure 24 depicts the nucleotide sequence of the construct designated Lys121-Val200; Asn425-Lys432 (SEQ ID NO:21).

Figure 25 depicts the nucleotide sequence of the construct designated Val120-Ile201; Ile424-Ala433 (SEQ ID NO:22).

Figure 26 depicts the nucleotide sequence of the construct designated Val120-Ile201B; Ile424-Ala433 (SEQ ID NO:23).

Figure 27 depicts the nucleotide sequence of the construct designated Val120-Thr202; 20 Ile424-Ala433 (SEQ ID NO:24).

Figure 28 depicts the nucleotide sequence of the construct designated Val127-Asn195 (SEQ ID NO:25).

Figure 29 depicts the nucleotide sequence of the construct designated Val127-Asn195; Arg426-Gly431 (SEQ ID NO:26).

25

30

5

10

15

Detailed Description of the Invention

The practice of the present invention will employ, unless otherwise indicated, conventional methods of protein chemistry, viral immunobiology, molecular biology and recombinant DNA techniques within the skill of the art. Such techniques are explained fully in the literature. See, e.g., T.E. Creighton, <u>Proteins: Structures and Molecular Properties</u> (W.H. Freeman and Company, 1993); Nelson L.M. and Jerome H.K. <u>HIV Protocols</u> in Methods in Molecular Medicine, vol. 17, 1999; Sambrook, et al., <u>Molecular Cloning: A</u>

<u>Laboratory Manual</u> (Cold Spring Harbor Laboratory, 1989); F.M. Ausubel et al. <u>Current Protocols in Molecular Biology</u>, Greene Publishing Associates & Wiley Interscience New York; and Lipkowitz and Boyd, <u>Reviews in Computational Chemistry</u>, volumes 1-present (Wiley-VCH, New York, New York, 1999).

It must be noted that, as used in this specification and the appended claims, the singular forms "a", "an" and "the" include plural referents unless the content clearly dictates otherwise. Thus, for example, reference to "a polypeptide" includes a mixture of two or more polypeptides, and the like.

10 Definitions

5

15

20

25

30

In describing the present invention, the following terms will be employed, and are intended to be defined as indicated below.

The terms "polypeptide," and "protein" are used interchangeably herein to denote any polymer of amino acid residues. The terms encompass peptides, oligopeptides, dimers, multimers, and the like. Such polypeptides can be derived from natural sources or can be synthesized or recombinantly produced. The terms also include postexpression modifications of the polypeptide, for example, glycosylation, acetylation, phosphorylation, etc.

A polypeptide as defined herein is generally made up of the 20 natural amino acids Ala (A), Arg (R), Asn (N), Asp (D), Cys (C), Gln (Q), Glu (E), Gly (G), His (H), Ile (I), Leu (L), Lys (K), Met (M), Phe (F), Pro (P), Ser (S), Thr (T), Trp (W), Tyr (Y) and Val (V) and may also include any of the several known amino acid analogs, both naturally occurring and synthesized analogs, such as but not limited to homoisoleucine, asaleucine, 2- (methylenecyclopropyl)glycine, S-methylcysteine, S-(prop-l-enyl)cysteine, homoserine, omithine, norleucine, norvaline, homoarginine, 3-(3-carboxyphenyl)alanine, cyclohexylalanine, mimosine, pipecolic acid, 4-methylglutamic acid, canavanine, 2,3-diaminopropionic acid, and the like. Further examples of polypeptide agents which will find use in the present invention are set forth below.

By "geometry" or "tertiary structure" of a polypeptide or protein is meant the overall 3-D configuration of the protein. As described herein, the geometry can be determined, for example, by crystallography studies or by using various programs or algorithms which predict the geometry based on interactions between the amino acids making up the primary and secondary structures.

By "wild type" polypeptide, polypeptide agent or polypeptide drug, is meant a naturally occurring polypeptide sequence, and its corresponding secondary structure. An "isolated" or "purified" protein or polypeptide is a protein which is separate and discrete from a whole organism with which the protein is normally associated in nature. It is apparent that the term denotes proteins of various levels of purity. Typically, a composition containing a purified protein will be one in which at least about 35%, preferably at least about 40-50%, more preferably, at least about 75-85%, and most preferably at least about 90% or more, of the total protein in the composition will be the protein in question.

5

10

15

20

25

30

By "Env polypeptide" is meant a molecule derived from an envelope protein, preferably from HIV Env. The envelope protein of HIV-1 is a glycoprotein of about 160 kd (gp160). During virus infection of the host cell, gp160 is cleaved by host cell proteases to form gp120 and the integral membrane protein, gp41. The gp41 portion is anchored in (and spans) the membrane bilayer of virion, while the gp120 segment protrudes into the surrounding environment. As there is no covalent attachment between gp120 and gp41, free gp120 is released from the surface of virions and infected cells. Env polypeptides may also include gp140 polypeptides. Env polypeptides can exist as monomers, dimers or multimers.

By a "gp120 polypeptide" is meant a molecule derived from a gp120 region of the Env polypeptide. Preferably, the gp120 polypeptide is derived from HIV Env. The primary amino acid sequence of gp120 is approximately 511 amino acids, with a polypeptide core of about 60,000 daltons. The polypeptide is extensively modified by N-linked glycosylation to increase the apparent molecular weight of the molecule to 120,000 daltons. The amino acid sequence of gp120 contains five relatively conserved domains interspersed with five hypervariable domains. The positions of the 18 cysteine residues in the gp120 primary sequence of the HIV-1_{HXB-2} (hereinafter "HXB-2") strain, and the positions of 13 of the approximately 24 N-linked glycosylation sites in the gp120 sequence are common to most, if not all, gp120 sequences. The hypervariable domains contain extensive amino acid substitutions, insertions and deletions. Despite this variation, most, if not all, gp120 sequences preserve the virus's ability to bind to the viral receptor CD4. A "gp120 polypeptide" includes both single subunits or multimers.

Env polypeptides (e.g., gp120, gp140 and gp160) include a "bridging sheet" comprised of 4 anti-parallel β -strands (β -2, β -3, β -20 and β -21) that form a β -sheet. Extruding from one pair of the β -strands (β -2 and β -3) are two loops, V1 and V2. The β -2

sheet occurs at approximately amino acid residue 119 (Cys) to amino acid residue 123 (Thr) while β-3 occurs at approximately amino acid residue 199 (Ser) to amino acid residue 201 (Ile), relative to HXB-2. The "V1/V2 region" occurs at approximately amino acid positions 126 (Cys) to residue 196 (Cys), relative to HXB-2. (see, e.g., Wyatt et al. (1995) *J. Virol.* 69:5723-5733; Stamatatos et al. (1998) *J. Virol.* 72:7840-7845). Extruding from the second pair of β-strands (β-20 and β-21) is a "small-loop" structure, also referred to herein as "the bridging sheet small loop." In HXB-2, β-20 extends from about amino acid residue 422 (Gln) to amino acid residue 426 (Met) while β-21 extends from about amino acid residue 430 (Val) to amino acid residue 435 (Tyr). In variant SF162, the Met-426 is an Arg (R) residue. The "small loop" extends from about amino acid residue 427 (Trp) through 429 (Lys), relative to HXB-2. A representative diagram of gp120 showing the bridging sheet, the small loop, and V1/V2 is shown in Figure 1. In addition, alignment of the amino acid sequences of Env polypeptide gp160 of selected variants is shown, relative to HXB-2, in Figures 2A-C.

5

10

15

20

25

30

Furthermore, an "Env polypeptide" or "gp120 polypeptide" as defined herein is not limited to a polypeptide having the exact sequence described herein. Indeed, the HIV genome is in a state of constant flux and contains several variable domains which exhibit relatively high degrees of variability between isolates. It is readily apparent that the terms encompass Env (e.g., gp120) polypeptides from any of the identified HIV isolates, as well as newly identified isolates, and subtypes of these isolates. Descriptions of structural features are given herein with reference to HXB-2. One of ordinary skill in the art in view of the teachings of the present disclosure and the art can determine corresponding regions in other HIV variants (e.g., isolates HIV_{IIIb}, HIV_{SF2}, HIV-1_{SF162}, HIV-1_{SF170}, HIV_{LAV}, HIV_{LAV}, HIV_{MN}, HIV-1_{CM235},, HIV-1_{US4}, other HIV-1 strains from diverse subtypes(e.g., subtypes, A through G, and O), HIV-2 strains and diverse subtypes (e.g., HIV-2_{UC1} and HIV-2_{UC2}), and simian immunodeficiency virus (SIV). (See, e.g., Virology, 3rd Edition (W.K. Joklik ed. 1988); Fundamental Virology, 2nd Edition (B.N. Fields and D.M. Knipe, eds. 1991); Virology, 3rd Edition (Fields, BN, DM Knipe, PM Howley, Editors, 1996, Lippincott-Raven, Philadelphia, PA; for a description of these and other related viruses), using for example, sequence comparison programs (e.g., BLAST and others described herein) or identification and alignment of structural features (e.g., a program such as the "ALB" program described herein that can identify \beta-sheet regions). The actual amino acid sequences of the modified Env polypeptides can be based on any HIV variant.

Additionally, the term "Env polypeptide" (e.g., "gp120 polypeptide") encompasses proteins which include additional modifications to the native sequence, such as additional internal deletions, additions and substitutions. These modifications may be deliberate, as through site-directed mutagenesis, or may be accidental, such as through naturally occurring mutational events. Thus, for example, if the Env polypeptide is to be used in vaccine compositions, the modifications must be such that immunological activity (i.e., the ability to elicit an antibody response to the polypeptide) is not lost. Similarly, if the polypeptides are to be used for diagnostic purposes, such capability must be retained.

5

10

15

20

25

30

Thus, a "modified Env polypeptide" is an Env polypeptide (e.g., gp120 as defined above), which has been manipulated to delete or replace all or a part of the bridging sheet portion and, optionally, the variable regions V1 and V2. Generally, modified Env (e.g., gp120) polypeptides have enough of the bridging sheet removed to expose the CD4 binding site, but leave enough of the structure to allow correct folding (e.g., correct geometry). Thus, modifications to the β -20 and β -21 regions (between about amino acid residues 420 and 435 relative to HXB-2) are preferred. Additionally, modifications to the β -2 and β -3 regions (between about amino acid residues 119 (Cys) and 201 (Ile)) and modifications (e.g., truncations) to the V1 and V2 loop regions may also be made. Although not all possible β -sheet and V1/V2 modifications have been exemplified herein, it is to be understood that other disrupting modifications are also encompassed by the present invention.

Normally, such a modified polypeptide is capable of secretion into growth medium in which an organism expressing the protein is cultured. However, for purposes of the present invention, such polypeptides may also be recovered intracellularly. Secretion into growth media is readily determined using a number of detection techniques, including, e.g., polyacrylamide gel electrophoresis and the like, and immunological techniques such as Western blotting and immunoprecipitation assays as described in, e.g., International Publication No. WO 96/04301, published February 15, 1996.

A gp120 or other Env polypeptide is produced "intracellularly" when it is found within the cell, either associated with components of the cell, such as in association with the endoplasmic reticulum (ER) or the Golgi Apparatus, or when it is present in the soluble cellular fraction. The gp120 and other Env polypeptides of the present invention may also be secreted into growth medium so long as sufficient amounts of the polypeptides remain

present within the cell such that they can be purified from cell lysates using techniques described herein.

5

10

15

20

25

30

An "immunogenic" gp120 or other Env protein is a molecule that includes at least one epitope such that the molecule is capable of either eliciting an immunological reaction in an individual to which the protein is administered or, in the diagnostic context, is capable of reacting with antibodies directed against the HIV in question.

By "epitope" is meant a site on an antigen to which specific B cells and/or T cells respond, rendering the molecule including such an epitope capable of eliciting an immunological reaction or capable of reacting with HIV antibodies present in a biological sample. The term is also used interchangeably with "antigenic determinant" or "antigenic determinant site." An epitope can comprise 3 or more amino acids in a spatial conformation unique to the epitope. Generally, an epitope consists of at least 5 such amino acids and, more usually, consists of at least 8-10 such amino acids. Methods of determining spatial conformation of amino acids are known in the art and include, for example, x-ray crystallography and 2-dimensional nuclear magnetic resonance. Furthermore, the identification of epitopes in a given protein is readily accomplished using techniques well known in the art, such as by the use of hydrophobicity studies and by site-directed serology. See, also, Geysen et al., Proc. Natl. Acad. Sci. USA (1984) 81:3998-4002 (general method of rapidly synthesizing peptides to determine the location of immunogenic epitopes in a given antigen); U.S. Patent No. 4,708,871 (procedures for identifying and chemically synthesizing epitopes of antigens); and Geysen et al., Molecular Immunology (1986) 23:709-715 (technique for identifying peptides with high affinity for a given antibody). Antibodies that recognize the same epitope can be identified in a simple immunoassay showing the ability of one antibody to block the binding of another antibody to a target antigen.

An "immunological response" or "immune response" as used herein is the development in the subject of a humoral and/or a cellular immune response to the Env (e.g., gp120) polypeptide when the polypeptide is present in a vaccine composition. These antibodies may also neutralize infectivity, and/or mediate antibody-complement or antibody dependent cell cytotoxicity to provide protection to an immunized host. Immunological reactivity may be determined in standard immunoassays, such as a competition assays, well known in the art.

Techniques for determining amino acid sequence "similarity" are well known in the art. In general, "similarity" means the exact amino acid to amino acid comparison of two or more polypeptides at the appropriate place, where amino acids are identical or possess similar chemical and/or physical properties such as charge or hydrophobicity. A so-termed "percent similarity" then can be determined between the compared polypeptide sequences.

Techniques for determining nucleic acid and amino acid sequence identity also are well known in the art and include determining the nucleotide sequence of the mRNA for that gene (usually via a cDNA intermediate) and determining the amino acid sequence encoded thereby, and comparing this to a second amino acid sequence. In general, "identity" refers to an exact nucleotide to nucleotide or amino acid to amino acid correspondence of two polynucleotides or polypeptide sequences, respectively.

5

10

15

20

25

30

Two or more polynucleotide sequences can be compared by determining their "percent identity." Two or more amino acid sequences likewise can be compared by determining their "percent identity." The percent identity of two sequences, whether nucleic acid or peptide sequences, is generally described as the number of exact matches between two aligned sequences divided by the length of the shorter sequence and multiplied by 100. An approximate alignment for nucleic acid sequences is provided by the local homology algorithm of Smith and Waterman, Advances in Applied Mathematics 2:482-489 (1981). This algorithm can be extended to use with peptide sequences using the scoring matrix developed by Dayhoff, Atlas of Protein Sequences and Structure, M.O. Dayhoff ed., 5 suppl. 3:353-358, National Biomedical Research Foundation, Washington, D.C., USA, and normalized by Gribskov, Nucl. Acids Res. 14(6):6745-6763 (1986). An implementation of this algorithm for nucleic acid and peptide sequences is provided by the Genetics Computer Group (Madison, WI) in their BestFit utility application. The default parameters for this method are described in the Wisconsin Sequence Analysis Package Program Manual, Version 8 (1995) (available from Genetics Computer Group, Madison, WI). Other equally suitable programs for calculating the percent identity or similarity between sequences are generally known in the art.

For example, percent identity of a particular nucleotide sequence to a reference sequence can be determined using the homology algorithm of Smith and Waterman with a default scoring table and a gap penalty of six nucleotide positions. Another method of establishing percent identity in the context of the present invention is to use the MPSRCH

package of programs copyrighted by the University of Edinburgh, developed by John F. Collins and Shane S. Sturrok, and distributed by IntelliGenetics, Inc. (Mountain View, CA). From this suite of packages, the Smith-Waterman algorithm can be employed where default parameters are used for the scoring table (for example, gap open penalty of 12, gap extension penalty of one, and a gap of six). From the data generated, the "Match" value reflects "sequence identity." Other suitable programs for calculating the percent identity or similarity between sequences are generally known in the art, such as the alignment program BLAST, which can also be used with default parameters. For example, BLASTN and BLASTP can be used with the following default parameters: genetic code = standard; filter = none; strand = both; cutoff = 60; expect = 10; Matrix = BLOSUM62; Descriptions = 50 sequences; sort by = HIGH SCORE; Databases = non-redundant, GenBank + EMBL + DDBJ + PDB + GenBank CDS translations + Swiss protein + Spupdate + PIR. Details of these programs can be found at the following internet address: http://www.ncbi.nlm.gov/cgi-bin/BLAST.

One of skill in the art can readily determine the proper search parameters to use for a given sequence in the above programs. For example, the search parameters may vary based on the size of the sequence in question. Thus, for example, a representative embodiment of the present invention would include an isolated polynucleotide having X contiguous nucleotides, wherein (i) the X contiguous nucleotides have at least about 50% identity to Y contiguous nucleotides derived from any of the sequences described herein, (ii) X equals Y, and (iii) X is greater than or equal to 6 nucleotides and up to 5000 nucleotides, preferably greater than or equal to 8 nucleotides and up to 5000 nucleotides, more preferably 10-12 nucleotides and up to 5000 nucleotides, and even more preferably 15-20 nucleotides, up to the number of nucleotides present in the full-length sequences described herein (e.g., see the Sequence Listing and claims), including all integer values falling within the above-described ranges.

The synthetic expression cassettes (and purified polynucleotides) of the present invention include related polynucleotide sequences having about 80% to 100%, greater than 80-85%, preferably greater than 90-92%, more preferably greater than 95%, and most preferably greater than 98% sequence (including all integer values falling within these described ranges) identity to the synthetic expression cassette sequences disclosed herein (for example, to the claimed sequences or other sequences of the present invention) when the sequences of the present invention are used as the query sequence.

Computer programs are also available to determine the likelihood of certain polypeptides to form structures such as β-sheets. One such program, described herein, is the "ALB" program for protein and polypeptide secondary structure calculation and predication. In addition, secondary protein structure can be predicted from the primary amino acid sequence, for example using protein crystal structure and aligning the protein sequence related to the crystal structure (e.g., using Molecular Operating Environment (MOE) programs available from the Chemical Computing Group Inc., Montreal, P.Q., Canada). Other methods of predicting secondary structures are described, for example, in Garnier et al. (1996) Methods Enzymol. 266:540-553; Geourjon et al. (1995) Comput. Applic. Biosci. 11:681-684; Levin (1997) Protein Eng. 10:771-776; and Rost et al. (1993) J. Molec. Biol. 232:584-599.

Homology can also be determined by hybridization of polynucleotides under conditions which form stable duplexes between homologous regions, followed by digestion with single-stranded-specific nuclease(s), and size determination of the digested fragments. Two DNA, or two polypeptide sequences are "substantially homologous" to each other when the sequences exhibit at least about 80%-85%, preferably at least about 90%, and most preferably at least about 95%-98% sequence identity over a defined length of the molecules, as determined using the methods above. As used herein, substantially homologous also refers to sequences showing complete identity to the specified DNA or polypeptide sequence. DNA sequences that are substantially homologous can be identified in a Southern hybridization experiment under, for example, stringent conditions, as defined for that particular system. Defining appropriate hybridization conditions is within the skill of the art. See, e.g., Sambrook et al., supra; DNA Cloning, supra; Nucleic Acid Hybridization, supra.

A "coding sequence" or a sequence which "encodes" a selected protein, is a nucleic acid sequence which is transcribed (in the case of DNA) and translated (in the case of mRNA) into a polypeptide *in vitro* or *in vivo* when placed under the control of appropriate regulatory sequences. The boundaries of the coding sequence are determined by a start codon at the 5' (amino) terminus and a translation stop codon at the 3' (carboxy) terminus. A coding sequence can include, but is not limited to cDNA from viral nucleotide sequences as well as synthetic and semisynthetic DNA sequences and sequences including base analogs. A transcription termination sequence may be located 3' to the coding sequence.

"Control elements" refers collectively to promoter sequences, ribosome binding sites, polyadenylation signals, transcription termination sequences, upstream regulatory domains, enhancers, and the like, which collectively provide for the transcription and translation of a coding sequence in a host cell. Not all of these control elements need always be present so long as the desired gene is capable of being transcribed and translated.

5

10

15

20

25

30

A control element "directs the transcription" of a coding sequence in a cell when RNA polymerase will bind the promoter sequence and transcribe the coding sequence into mRNA, which is then translated into the polypeptide encoded by the coding sequence.

"Operably linked" refers to an arrangement of elements wherein the components so described are configured so as to perform their usual function. Thus, control elements operably linked to a coding sequence are capable of effecting the expression of the coding sequence when RNA polymerase is present. The control elements need not be contiguous with the coding sequence, so long as they function to direct the expression thereof. Thus, for example, intervening untranslated yet transcribed sequences can be present between, e.g., a promoter sequence and the coding sequence and the promoter sequence can still be considered "operably linked" to the coding sequence.

"Recombinant" as used herein to describe a nucleic acid molecule means a polynucleotide of genomic, cDNA, semisynthetic, or synthetic origin which, by virtue of its origin or manipulation: (1) is not associated with all or a portion of the polynucleotide with which it is associated in nature; and/or (2) is linked to a polynucleotide other than that to which it is linked in nature. The term "recombinant" as used with respect to a protein or polypeptide means a polypeptide produced by expression of a recombinant polynucleotide. "Recombinant host cells," "host cells," "cells," "cell lines," "cell cultures," and other such terms denoting procaryotic microorganisms or eucaryotic cell lines cultured as unicellular entities, are used interchangeably, and refer to cells which can be, or have been, used as recipients for recombinant vectors or other transfer DNA, and include the progeny of the original cell which has been transfected. It is understood that the progeny of a single parental cell may not necessarily be completely identical in morphology or in genomic or total DNA complement to the original parent, due to accidental or deliberate mutation. Progeny of the parental cell which are sufficiently similar to the parent to be characterized by the relevant property, such as the presence of a nucleotide sequence encoding a desired peptide, are included in the progeny intended by this definition, and are covered by the above terms.

By "vertebrate subject" is meant any member of the subphylum chordata, including, without limitation, humans and other primates, including non-human primates such as chimpanzees and other apes and monkey species; farm animals such as cattle, sheep, pigs, goats and horses; domestic mammals such as dogs and cats; laboratory animals including rodents such as mice, rats and guinea pigs; birds, including domestic, wild and game birds such as chickens, turkeys and other gallinaceous birds, ducks, geese, and the like. The term does not denote a particular age. Thus, both adult and newborn individuals are intended to be covered.

As used herein, a "biological sample" refers to a sample of tissue or fluid isolated from an individual, including but not limited to, for example, blood, plasma, serum, fecal matter, urine, bone marrow, bile, spinal fluid, lymph fluid, samples of the skin, external secretions of the skin, respiratory, intestinal, and genitourinary tracts, samples derived from the gastric epithelium and gastric mucosa, tears, saliva, milk, blood cells, organs, biopsies and also samples of *in vitro* cell culture constituents including but not limited to conditioned media resulting from the growth of cells and tissues in culture medium, e.g., recombinant cells, and cell components.

The terms "label" and "detectable label" refer to a molecule capable of detection, including, but not limited to, radioactive isotopes, fluorescers, chemiluminescers, enzymes, enzyme substrates, enzyme cofactors, enzyme inhibitors, chromophores, dyes, metal ions, metal sols, ligands (e.g., biotin or haptens) and the like. The term "fluorescer" refers to a substance or a portion thereof which is capable of exhibiting fluorescence in the detectable range. Particular examples of labels which may be used with the invention include, but are not limited to fluorescein, rhodamine, dansyl, umbelliferone, Texas red, luminol, acradimum esters, NADPH, α - β -galactosidase, horseradish peroxidase, glucose oxidase, alkaline phosphatase and urease.

Overview

5

10

15

20

25

30

The present invention concerns modified Env polypeptide molecules (e.g., glycoprotein ("gp") 120). Without being bound by a particular theory, it appears that it has been difficult to generate immunological responses against Env because the CD4 binding site is buried between the outer domain, the inner domain and the V1/V2 domains. Thus, although deletion of the V1/V2 domain may render the virus more susceptible to

neutralization by monoclonal antibody directed to the CD4 site, the bridging sheet covering most of the CD4 binding domain may prevent an antibody response. Thus, the present invention provides Env polypeptides that maintain their general overall structure yet expose the CD4 binding domain. This allows the generation of an immune response (e.g., an antibody response) to epitopes in or near the CD4 binding site.

Various forms of the different embodiments of the invention, described herein, may be combined.

β-Sheet Conformations

5

10

15

20

25

30

In the present invention, location of the β -sheet structures were identified relative to 3-D (crystal) structure of an HXB-2 crystallized Env protein (see, Example 1A). Based on this structure, constructs encoding polypeptides having replacements and or excisions which maintain overall geometry while exposing the CD4 binding site were designed. In particular, the crystal structure of HXB-2 was downloaded from the Brookhaven Database. Using the default parameters of the Loop Search feature of the Biopolymer module of the Sybyl molecular modeling package, homology and fit of amino acids which could replace the native loops between β -strands yet maintain overall tertiary structure were determined. Constructs encoding the modified Env polypeptides were then designed (Example 1.B.).

Thus, the modified Env polypeptides typically have enough of the bridging sheet removed to expose the CD4 groove, but have enough of the structure to allow correct folding of the Env glycoprotein. Exemplary constructs are described below.

Polypeptide Production

The polypeptides of the present invention can be produced in any number of ways which are well known in the art.

In one embodiment, the polypeptides are generated using recombinant techniques, well known in the art. In this regard, oligonucleotide probes can be devised based on the known sequences of the Env (e.g., gp120) polypeptide genome and used to probe genomic or cDNA libraries for Env genes. The gene can then be further isolated using standard techniques and, e.g., restriction enzymes employed to truncate the gene at desired portions of the full-length sequence. Similarly, the Env gene(s) can be isolated directly from cells and tissues containing the same, using known techniques, such as phenol extraction and the

sequence further manipulated to produce the desired truncations. See, e.g., Sambrook et al., supra, for a description of techniques used to obtain and isolate DNA.

The genes encoding the modified (e.g., truncated and/or substituted) polypeptides can be produced synthetically, based on the known sequences. The nucleotide sequence can be designed with the appropriate codons for the particular amino acid sequence desired. The complete sequence is generally assembled from overlapping oligonucleotides prepared by standard methods and assembled into a complete coding sequence. See, e.g., Edge (1981) Nature 292:756; Nambair et al. (1984) Science 223:1299; Jay et al. (1984) J. Biol. Chem. 259:6311; Stemmer et al. (1995) Gene 164:49-53.

10

15

5

Recombinant techniques are readily used to clone a gene encoding an Env polypeptide gene which can then be mutagenized *in vitro* by the replacement of the appropriate base pair(s) to result in the codon for the desired amino acid. Such a change can include as little as one base pair, effecting a change in a single amino acid, or can encompass several base pair changes. Alternatively, the mutations can be effected using a mismatched primer which hybridizes to the parent nucleotide sequence (generally cDNA corresponding to the RNA sequence), at a temperature below the melting temperature of the mismatched duplex. The primer can be made specific by keeping primer length and base composition within relatively narrow limits and by keeping the mutant base centrally located. See, *e.g.*, Innis et al, (1990) PCR Applications: Protocols for Functional Genomics; Zoller and Smith, *Methods Enzymol.* (1983) 100:468. Primer extension is effected using DNA polymerase, the product cloned and clones containing the mutated DNA, derived by segregation of the primer extended strand, selected. Selection can be accomplished using the mutant primer as a hybridization probe. The technique is also applicable for generating multiple point mutations. See, e.g., Dalbie-McFarland et al. *Proc. Natl. Acad. Sci USA* (1982) 79:6409.

25

30

20

Once coding sequences for the desired proteins have been isolated or synthesized, they can be cloned into any suitable vector or replicon for expression. As will be apparent from the teachings herein, a wide variety of vectors encoding modified polypeptides can be generated by creating expression constructs which operably link, in various combinations, polynucleotides encoding Env polypeptides having deletions or mutation therein. Thus, polynucleotides encoding a particular deleted V1/V2 region can be operably linked with polynucleotides encoding polypeptides having deletions or replacements in the small loop

region and the construct introduced into a host cell for polypeptide expression. Non-limiting examples of such combinations are discussed in the Examples.

Numerous cloning vectors are known to those of skill in the art, and the selection of an appropriate cloning vector is a matter of choice. Examples of recombinant DNA vectors for cloning and host cells which they can transform include the bacteriophage λ (E. coli), pBR322 (E. coli), pACYC177 (E. coli), pKT230 (gram-negative bacteria), pGV1106 (gram-negative bacteria), pLAFR1 (gram-negative bacteria), pME290 (non-E. coli gram-negative bacteria), pHV14 (E. coli and Bacillus subtilis), pBD9 (Bacillus), pIJ61 (Streptomyces), pUC6 (Streptomyces), YIp5 (Saccharomyces), YCp19 (Saccharomyces) and bovine papilloma virus (mammalian cells). See, generally, DNA Cloning: Vols. I & II, supra; Sambrook et al., supra; B. Perbal, supra.

5

10

15

20

25

30

Insect cell expression systems, such as baculovirus systems, can also be used and are known to those of skill in the art and described in, e.g., Summers and Smith, *Texas*Agricultural Experiment Station Bulletin No. 1555 (1987). Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, inter alia, Invitrogen, San Diego CA ("MaxBac" kit).

Plant expression systems can also be used to produce the modified Env proteins. Generally, such systems use virus-based vectors to transfect plant cells with heterologous genes. For a description of such systems see, e.g., Porta et al., *Mol. Biotech.* (1996) 5:209-221; and Hackland et al., *Arch. Virol.* (1994) 139:1-22.

Viral systems, such as a vaccinia based infection/transfection system, as described in Tomei et al., *J. Virol.* (1993) 67:4017-4026 and Selby et al., *J. Gen. Virol.* (1993) 74:1103-1113, will also find use with the present invention. In this system, cells are first transfected *in vitro* with a vaccinia virus recombinant that encodes the bacteriophage T7 RNA polymerase. This polymerase displays exquisite specificity in that it only transcribes templates bearing T7 promoters. Following infection, cells are transfected with the DNA of interest, driven by a T7 promoter. The polymerase expressed in the cytoplasm from the vaccinia virus recombinant transcribes the transfected DNA into RNA which is then translated into protein by the host translational machinery. The method provides for high level, transient, cytoplasmic production of large quantities of RNA and its translation product(s).

The gene can be placed under the control of a promoter, ribosome binding site (for bacterial expression) and, optionally, an operator (collectively referred to herein as "control" elements), so that the DNA sequence encoding the desired Env polypeptide is transcribed into RNA in the host cell transformed by a vector containing this expression construction. The coding sequence may or may not contain a signal peptide or leader sequence. With the present invention, both the naturally occurring signal peptides or heterologous sequences can be used. Leader sequences can be removed by the host in post-translational processing. See, e.g., U.S. Patent Nos. 4,431,739; 4,425,437; 4,338,397. Such sequences include, but are not limited to, the TPA leader, as well as the honey bee mellitin signal sequence.

5

10

15

20

25

30

Other regulatory sequences may also be desirable which allow for regulation of expression of the protein sequences relative to the growth of the host cell. Such regulatory sequences are known to those of skill in the art, and examples include those which cause the expression of a gene to be turned on or off in response to a chemical or physical stimulus, including the presence of a regulatory compound. Other types of regulatory elements may also be present in the vector, for example, enhancer sequences.

The control sequences and other regulatory sequences may be ligated to the coding sequence prior to insertion into a vector. Alternatively, the coding sequence can be cloned directly into an expression vector which already contains the control sequences and an appropriate restriction site.

In some cases it may be necessary to modify the coding sequence so that it may be attached to the control sequences with the appropriate orientation; i.e., to maintain the proper reading frame. Mutants or analogs may be prepared by the deletion of a portion of the sequence encoding the protein, by insertion of a sequence, and/or by substitution of one or more nucleotides within the sequence. Techniques for modifying nucleotide sequences, such as site-directed mutagenesis, are well known to those skilled in the art. See, e.g., Sambrook et al., supra; DNA Cloning, Vols. I and II, supra; Nucleic Acid Hybridization, supra.

The expression vector is then used to transform an appropriate host cell. A number of mammalian cell lines are known in the art and include immortalized cell lines available from the American Type Culture Collection (ATCC), such as, but not limited to, Chinese hamster ovary (CHO) cells, HeLa cells, baby hamster kidney (BHK) cells, monkey kidney cells (COS), human hepatocellular carcinoma cells (e.g., Hep G2), Vero293 cells, as well as others. Similarly, bacterial hosts such as *E. coli*, *Bacillus subtilis*, and *Streptococcus spp.*, will find

use with the present expression constructs. Yeast hosts useful in the present invention include inter alia, Saccharomyces cerevisiae, Candida albicans, Candida maltosa, Hansenula polymorpha, Kluyveromyces fragilis, Kluyveromyces lactis, Pichia guillerimondii, Pichia pastoris, Schizosaccharomyces pombe and Yarrowia lipolytica. Insect cells for use with baculovirus expression vectors include, inter alia, Aedes aegypti, Autographa californica, Bombyx mori, Drosophila melanogaster, Spodoptera frugiperda, and Trichoplusia ni.

5

10

15

20

25

30

Depending on the expression system and host selected, the proteins of the present invention are produced by growing host cells transformed by an expression vector described above under conditions whereby the protein of interest is expressed. The selection of the appropriate growth conditions is within the skill of the art.

In one embodiment, the transformed cells secrete the polypeptide product into the surrounding media. Certain regulatory sequences can be included in the vector to enhance secretion of the protein product, for example using a tissue plasminogen activator (TPA) leader sequence, a γ-interferon signal sequence or other signal peptide sequences from known secretory proteins. The secreted polypeptide product can then be isolated by various techniques described herein, for example, using standard purification techniques such as but not limited to, hydroxyapatite resins, column chromatography, ion-exchange chromatography, size-exclusion chromatography, electrophoresis, HPLC, immunoadsorbent techniques, affinity chromatography, immunoprecipitation, and the like..

Alternatively, the transformed cells are disrupted, using chemical, physical or mechanical means, which lyse the cells yet keep the Env polypeptides substantially intact. Intracellular proteins can also be obtained by removing components from the cell wall or membrane, e.g., by the use of detergents or organic solvents, such that leakage of the Env polypeptides occurs. Such methods are known to those of skill in the art and are described in, e.g., *Protein Purification Applications: A Practical Approach*, (E.L.V. Harris and S. Angal, Eds., 1990)

For example, methods of disrupting cells for use with the present invention include but are not limited to: sonication or ultrasonication; agitation; liquid or solid extrusion; heat treatment; freeze-thaw; desiccation; explosive decompression; osmotic shock; treatment with lytic enzymes including proteases such as trypsin, neuraminidase and lysozyme; alkali treatment; and the use of detergents and solvents such as bile salts, sodium dodecylsulphate,

Triton, NP40 and CHAPS. The particular technique used to disrupt the cells is largely a matter of choice and will depend on the cell type in which the polypeptide is expressed, culture conditions and any pre-treatment used.

Following disruption of the cells, cellular debris is removed, generally by centrifugation, and the intracellularly produced Env polypeptides are further purified, using standard purification techniques such as but not limited to, column chromatography, ion-exchange chromatography, size-exclusion chromatography, electrophoresis, HPLC, immunoadsorbent techniques, affinity chromatography, immunoprecipitation, and the like.

5

10

15

20

25

30

For example, one method for obtaining the intracellular Env polypeptides of the present invention involves affinity purification, such as by immunoaffinity chromatography using anti-Env specific antibodies, or by lectin affinity chromatography. Particularly preferred lectin resins are those that recognize mannose moieties such as but not limited to resins derived from *Galanthus nivalis* agglutinin (GNA), *Lens culinaris* agglutinin (LCA or lentil lectin), *Pisum sativum* agglutinin (PSA or pea lectin), *Narcissus pseudonarcissus* agglutinin (NPA) and *Allium ursinum* agglutinin (AUA). The choice of a suitable affinity resin is within the skill in the art. After affinity purification, the Env polypeptides can be further purified using conventional techniques well known in the art, such as by any of the techniques described above.

It may be desirable to produce Env (e.g., gp120) complexes, either with itself or other proteins. Such complexes are readily produced by e.g., co-transfecting host cells with constructs encoding for the Env (e.g., gp120) and/or other polypeptides of the desired complex. Co-transfection can be accomplished either in trans or cis, i.e., by using separate vectors or by using a single vector which bears both of the Env and other gene. If done using a single vector, both genes can be driven by a single set of control elements or, alternatively, the genes can be present on the vector in individual expression cassettes, driven by individual control elements. Following expression, the proteins will spontaneously associate.

Alternatively, the complexes can be formed by mixing the individual proteins together which have been produced separately, either in purified or semi-purified form, or even by mixing culture media in which host cells expressing the proteins, have been cultured. See, International Publication No. WO 96/04301, published February 15, 1996, for a description of such complexes.

5

10

15

20

25

30

Relatively small polypeptides, i.e., up to about 50 amino acids in length, can be conveniently synthesized chemically, for example by any of several techniques that are known to those skilled in the peptide art. In general, these methods employ the sequential addition of one or more amino acids to a growing peptide chain. Normally, either the amino or carboxyl group of the first amino acid is protected by a suitable protecting group. The protected or derivatized amino acid can then be either attached to an inert solid support or utilized in solution by adding the next amino acid in the sequence having the complementary (amino or carboxyl) group suitably protected, under conditions that allow for the formation of an amide linkage. The protecting group is then removed from the newly added amino acid residue and the next amino acid (suitably protected) is then added, and so forth. After the desired amino acids have been linked in the proper sequence, any remaining protecting groups (and any solid support, if solid phase synthesis techniques are used) are removed sequentially or concurrently, to render the final polypeptide. By simple modification of this general procedure, it is possible to add more than one amino acid at a time to a growing chain, for example, by coupling (under conditions which do not racemize chiral centers) a protected tripeptide with a properly protected dipeptide to form, after deprotection, a pentapeptide. See, e.g., J. M. Stewart and J. D. Young, Solid Phase Peptide Synthesis (Pierce Chemical Co., Rockford, IL 1984) and G. Barany and R. B. Merrifield, The Peptides: Analysis, Synthesis, Biology, editors E. Gross and J. Meienhofer, Vol. 2, (Academic Press, New York, 1980), pp. 3-254, for solid phase peptide synthesis techniques; and M. Bodansky, Principles of Peptide Synthesis, (Springer-Verlag, Berlin 1984) and E. Gross and J. Meienhofer, Eds., The Peptides: Analysis, Synthesis, Biology, Vol. 1, for classical solution synthesis.

Typical protecting groups include t-butyloxycarbonyl (Boc), 9-fluorenylmethoxycarbonyl (Fmoc) benzyloxycarbonyl (Cbz); p-toluenesulfonyl (Tx); 2,4-dinitrophenyl; benzyl (Bzl); biphenylisopropyloxycarboxy-carbonyl, t-amyloxycarbonyl, isobornyloxycarbonyl, o-bromobenzyloxycarbonyl, cyclohexyl, isopropyl, acetyl, o-nitrophenylsulfonyl and the like.

Typical solid supports are cross-linked polymeric supports. These can include divinylbenzene cross-linked-styrene-based polymers, for example, divinylbenzene-hydroxymethylstyrene copolymers, divinylbenzene-chloromethylstyrene copolymers and divinylbenzene-benzhydrylaminopolystyrene copolymers.

The polypeptide analogs of the present invention can also be chemically prepared by other methods such as by the method of simultaneous multiple peptide synthesis. See, e.g., Houghten *Proc. Natl. Acad. Sci. USA* (1985) 82:5131-5135; U.S. Patent No. 4,631,211.

Diagnostic and Vaccine Applications

5

10

15

20

25

30

The intracellularly produced Env polypeptides of the present invention, complexes thereof, or the polynucleotides coding therefor, can be used for a number of diagnostic and therapeutic purposes. For example, the proteins and polynucleotides or antibodies generated against the same, can be used in a variety of assays, to determine the presence of reactive antibodies/and or Env proteins in a biological sample to aid in the diagnosis of HIV infection or disease status or as measure of response to immunization.

The presence of antibodies reactive with the Env (e.g., gp120) polypeptides and, conversely, antigens reactive with antibodies generated thereto, can be detected using standard electrophoretic and immunodiagnostic techniques, including immunoassays such as competition, direct reaction, or sandwich type assays. Such assays include, but are not limited to, western blots; agglutination tests; enzyme-labeled and mediated immunoassays, such as ELISAs; biotin/avidin type assays; radioimmunoassays; immunoelectrophoresis; immunoprecipitation, etc. The reactions generally include revealing labels such as fluorescent, chemiluminescent, radioactive, or enzymatic labels or dye molecules, or other methods for detecting the formation of a complex between the antigen and the antibody or antibodies reacted therewith.

Solid supports can be used in the assays such as nitrocellulose, in membrane or microtiter well form; polyvinylchloride, in sheets or microtiter wells; polystyrene latex, in beads or microtiter plates; polyvinylidine fluoride; diazotized paper; nylon membranes; activated beads, and the like.

Typically, the solid support is first reacted with the biological sample (or the gp120 proteins), washed and then the antibodies, (or a sample suspected of containing antibodies), applied. After washing to remove any non-bound ligand, a secondary binder moiety is added under suitable binding conditions, such that the secondary binder is capable of associating selectively with the bound ligand. The presence of the secondary binder can then be detected using techniques well known in the art. Typically, the secondary binder will comprise an antibody directed against the antibody ligands. A number of anti-human immunoglobulin

(Ig) molecules are known in the art (e.g., commercially available goat anti-human Ig or rabbit anti-human Ig). Ig molecules for use herein will preferably be of the IgG or IgA type, however, IgM may also be appropriate in some instances. The Ig molecules can be readily conjugated to a detectable enzyme label, such as horseradish peroxidase, glucose oxidase, Beta-galactosidase, alkaline phosphatase and urease, among others, using methods known to those of skill in the art. An appropriate enzyme substrate is then used to generate a detectable signal.

5

10

15

20

25

30

Alternatively, a "two antibody sandwich" assay can be used to detect the proteins of the present invention. In this technique, the solid support is reacted first with one or more of the antibodies directed against Env (e.g., gp120), washed and then exposed to the test sample. Antibodies are again added and the reaction visualized using either a direct color reaction or using a labeled second antibody, such as an anti-immunoglobulin labeled with horseradish peroxidase, alkaline phosphatase or urease.

Assays can also be conducted in solution, such that the viral proteins and antibodies thereto form complexes under precipitating conditions. The precipitated complexes can then be separated from the test sample, for example, by centrifugation. The reaction mixture can be analyzed to determine the presence or absence of antibody-antigen complexes using any of a number of standard methods, such as those immunodiagnostic methods described above.

The modified Env proteins, produced as described above, or antibodies to the proteins, can be provided in kits, with suitable instructions and other necessary reagents, in order to conduct immunoassays as described above. The kit can also contain, depending on the particular immunoassay used, suitable labels and other packaged reagents and materials (i.e. wash buffers and the like). Standard immunoassays, such as those described above, can be conducted using these kits.

The Env polypeptides and polynucleotides encoding the polypeptides can also be used in vaccine compositions, individually or in combination, in e.g., prophylactic (i.e., to prevent infection) or therapeutic (to treat HIV following infection) vaccines. The vaccines can comprise mixtures of one or more of the modified Env proteins (or nucleotide sequences encoding the proteins), such as Env (e.g., gp120) proteins derived from more than one viral isolate. The vaccine may also be administered in conjunction with other antigens and immunoregulatory agents, for example, immunoglobulins, cytokines, lymphokines, and chemokines, including but not limited to IL-2, modified IL-2 (cys125-ser125), GM-CSF, IL-

12, γ-interferon, IP-10, MIP1β and RANTES. The vaccines may be administered as polypeptides or, alternatively, as naked nucleic acid vaccines (e.g., DNA), using viral vectors (e.g., retroviral vectors, adenoviral vectors, adeno-associated viral vectors) or non-viral vectors (e.g., liposomes, particles coated with nucleic acid or protein). The vaccines may also comprise a mixture of protein and nucleic acid, which in turn may be delivered using the same or different vehicles. The vaccine may be given more than once (e.g., a "prime" administration followed by one or more "boosts") to achieve the desired effects. The same composition can be administered as the prime and as the one or more boosts. Alternatively, different compositions can be used for priming and boosting.

5

10

15

20

25

30

The vaccines will generally include one or more "pharmaceutically acceptable excipients or vehicles" such as water, saline, glycerol, ethanol, etc. Additionally, auxiliary substances, such as wetting or emulsifying agents, pH buffering substances, and the like, may be present in such vehicles.

A carrier is optionally present which is a molecule that does not itself induce the production of antibodies harmful to the individual receiving the composition. Suitable carriers are typically large, slowly metabolized macromolecules such as proteins, polysaccharides, polylactic acids, polyglycollic acids, polymeric amino acids, amino acid copolymers, lipid aggregates (such as oil droplets or liposomes), and inactive virus particles. Such carriers are well known to those of ordinary skill in the art. Furthermore, the Env polypeptide may be conjugated to a bacterial toxoid, such as toxoid from diphtheria, tetanus, cholera, etc.

Adjuvants may also be used to enhance the effectiveness of the vaccines. Such adjuvants include, but are not limited to: (1) aluminum salts (alum), such as aluminum hydroxide, aluminum phosphate, aluminum sulfate, etc.; (2) oil-in-water emulsion formulations (with or without other specific immunostimulating agents such as muramyl peptides (see below) or bacterial cell wall components), such as for example (a) MF59 (International Publication No. WO 90/14837), containing 5% Squalene, 0.5% Tween 80, and 0.5% Span 85 (optionally containing various amounts of MTP-PE (see below), although not required) formulated into submicron particles using a microfluidizer such as Model 110Y microfluidizer (Microfluidics, Newton, MA), (b) SAF, containing 10% Squalane, 0.4% Tween 80, 5% pluronic-blocked polymer L121, and thr-MDP (see below) either microfluidized into a submicron emulsion or vortexed to generate a larger particle size

5

10

15

20

25

30

emulsion, and (c) RibiTM adjuvant system (RAS), (Ribi Immunochem, Hamilton, MT) containing 2% Squalene, 0.2% Tween 80, and one or more bacterial cell wall components from the group consisting of monophosphorylipid A (MPL), trehalose dimycolate (TDM), and cell wall skeleton (CWS), preferably MPL + CWS (DetoxTM); (3) saponin adjuvants, such as Stimulon™ (Cambridge Bioscience, Worcester, MA) may be used or particle generated therefrom such as ISCOMs (immunostimulating complexes); (4) Complete Freunds Adjuvant (CFA) and Incomplete Freunds Adjuvant (IFA); (5) cytokines, such as interleukins (IL-1, IL-2, etc.), macrophage colony stimulating factor (M-CSF), tumor necrosis factor (TNF), etc.; (6) detoxified mutants of a bacterial ADP-ribosylating toxin such as a cholera toxin (CT), a pertussis toxin (PT), or an E. coli heat-labile toxin (LT), particularly LT-K63 (where lysine is substituted for the wild-type amino acid at position 63) LT-R72 (where arginine is substituted for the wild-type amino acid at position 72), CT-S109 (where serine is substituted for the wild-type amino acid at position 109), and PT-K9/G129 (where lysine is substituted for the wild-type amino acid at position 9 and glycine substituted at position 129) (see, e.g., International Publication Nos. W093/13202 and W092/19265); and (7) other substances that act as immunostimulating agents to enhance the effectiveness of the composition.

Muramyl peptides include, but are not limited to, N-acetyl-muramyl-L-threonyl-D-isoglutamine (thr-MDP), N-acteyl-normuramyl-L-alanyl-D-isogluatme (nor-MDP), N-acetylmuramyl-L-alanyl-D-isogluatminyl-L-alanine-2-(l'-2'-dipalmitoyl-sn-glycero-3-huydroxyphosphoryloxy)-ethylamine (MTP-PE), etc.

Typically, the vaccine compositions are prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid vehicles prior to injection may also be prepared. The preparation also may be emulsified or encapsulated in liposomes for enhanced adjuvant effect, as discussed above.

The vaccines will comprise a therapeutically effective amount of the modified Env proteins, or complexes of the proteins, or nucleotide sequences encoding the same, and any other of the above-mentioned components, as needed. By "therapeutically effective amount" is meant an amount of a modified Env (e.g., gp120) protein which will induce a protective immunological response in the uninfected, infected or unexposed individual to which it is administered. Such a response will generally result in the development in the subject of a secretory, cellular and/or antibody-mediated immune response to the vaccine. Usually, such

a response includes but is not limited to one or more of the following effects; the production of antibodies from any of the immunological classes, such as immunoglobulins A, D, E, G or M; the proliferation of B and T lymphocytes; the provision of activation, growth and differentiation signals to immunological cells; expansion of helper T cell, suppressor T cell, and/or cytotoxic T cell.

5

10

15

20

25

30

Preferably, the effective amount is sufficient to bring about treatment or prevention of disease symptoms. The exact amount necessary will vary depending on the subject being treated; the age and general condition of the individual to be treated; the capacity of the individual's immune system to synthesize antibodies; the degree of protection desired; the severity of the condition being treated; the particular Env polypeptide selected and its mode of administration, among other factors. An appropriate effective amount can be readily determined by one of skill in the art. A "therapeutically effective amount" will fall in a relatively broad range that can be determined through routine trials.

Once formulated, the nucleic acid vaccines may be accomplished with or without viral vectors, as described above, by injection using either a conventional syringe or a gene gun, such as the Accell® gene delivery system (PowderJect Technologies, Inc., Oxford, England). Delivery of DNA into cells of the epidermis is particularly preferred as this mode of administration provides access to skin-associated lymphoid cells and provides for a transient presence of DNA in the recipient. Both nucleic acids and/or peptides can be injected either subcutaneously, epidermally, intradermally, intramucosally such as nasally, rectally and vaginally, intraperitoneally, intravenously, orally or intramuscularly. Other modes of administration include oral and pulmonary administration, suppositories, needle-less injection, transcutaneous and transdermal applications. Dosage treatment may be a single dose schedule or a multiple dose schedule. Administration of nucleic acids may also be combined with administration of peptides or other substances.

While the invention has been described in conjunction with the preferred specific embodiments thereof, it is to be understood that the foregoing description as well as the examples which follow are intended to illustrate and not limit the scope of the invention. Other aspects, advantages and modifications within the scope of the invention will be apparent to those skilled in the art to which the invention pertains.

Experimental

5

10

15

20

25

30

Below are examples of specific embodiments for carrying out the present invention. The examples are offered for illustrative purposes only, and are not intended to limit the scope of the present invention in any way.

Efforts have been made to ensure accuracy with respect to numbers used (e.g., amounts, temperatures, etc.), but some experimental error and deviation should, of course, be allowed for.

EXAMPLE 1

A.1. Best-Fit and Homology Searches

The crystal structure of HXB-2 gp 120 was downloaded from the Brookhaven database (COMPLEX (HIV ENVELOPE PROTEIN/CD4/FAB) 15-JUN-98 1GC1 TITLE: HIV-1 GP120 CORE COMPLEXED WITH CD4 AND A NEUTRALIZING HUMAN ANTIBODY). Beta strands 3, 2, 21, and 20 of gp 120 form a sheet near the CD4 binding site. Strands β -3 and β -2 are connected by the V1/V2 loop. Strands β -21 and β -20 are connected by another small loop. The H-bonds at the interface between strands β -2 and β -21 are the only connection between domains of the "lower" half of the protein (joining helix alpha 1 to the CD4 binding site). This beta sheet and these loops mask some antigens (e.g., antigens which may generate neutralizing antibodies) that are only exposed during the CD4 binding.

Constructs that remove enough of the beta sheet to expose the antigens in the CD4 binding site, but leave enough of the protein to allow correct folding were designed. Specifically targeted were modifications to the small loop and, optional deletion of the V1/V2 loops. Three different types of constructs were designed: (1) constructs encoding polypeptides that leave the number of residues making up the entire 4-strand beta sheet intact, but replace one or more residues; (2) constructs that encode polypeptide having at least one residue of at least one beta strand excised or (3) constructs encoding polypeptides having at least two residues of at least one beta strand excised. Thus, a total of 6 different turns were needed to rejoin the ends of the strands.

Initially, residues in the small loop (residues 427-430, relative to HXB-2) and connected beta strands (β -20 and β -21) were modified to contain Gly and Pro (common in beta turns). These sequences were then used as the target to match in each search. The

geometry of the target was matched to known proteins in the Brookhaven Protein Data Bank. In particular, 5-residue turns (including an overlapping single residue at the N-terminal, the 2 residue target turn and 2 overlapping residues at the C-terminal) were searched in the databases. In other words, these modified loops add a 2 residue turn that should be able to support a geometry that will maintain the beta-sheet structure of the wild type protein. The calculations were performed using the default parameters in the Loop Search feature of the Biopolymer module of the Syby1 molecular modeling package. In each case, the 25 best fits based on geometry alone were reviewed and, of those, several selected for homology and fit.

In addition, it was also determined what modifications could be made to remove most of the V1/V2 loop (residues 124-198, relative to HXB-2) yet leave the geometry of the protein intact. As with the small loop, constructs were also designed which excised one or more residues from the β -2 strand (residues 119-123 of HXB-2), the β -3 strand (residues 199-201 of HXB-2) or both β -2 and β -3. For these constructs, known loops were searched to match the geometry of a pentamer (including two remaining residues from the N-terminal side, a 2 residue turn and 1 C-terminal residue). For these searches, Gly-Gly was preferred as the insert along with at least one C-terminal substitution.

A.2. Small Loop Replacements

In one aspect, the native sequence was replaced with residues that expose the CD4 binding site, but leave the overall geometry of the protein relatively unchanged. For the small loop replacements, the target to match was: ASN425-MET426-GLY427-GLY428-GLY431. Results of the search are summarized in Table 1.

Table 1: Search of Small Loop (Asn425 through Gly431)

_	_
7	•
Z.	.)

20

5

10

15

Rank	Sequence	RMSD	% Homology	Seq Id No.
Best fit	LYS-ASP-SER-ASN-ASN	0.16689	62.5	27
3	TYR-GLY-LEU-GLY-LEU	0.220308	62.5	28
4	GLU-ARG-GLU-ASP-GLY	0.241754	62.5	29
7	ARG-LYS-GLY-GLY-ASN	0.24881	100	30
12	TRP-THR-GLY-SER-TYR	0.26417	83.33	31

30

Based on these results, constructs encoding Gly-Gly (#7), Gly-Ser (#12) or Gly-Gly-Asn (#7) were recommended.

As V1/V2 and one or more residues of β -2 and β -3 are also optionally deleted in the modified polypeptides of the invention, known loops to match the geometry of the V1/V2 loop were also searched. The V1/V2 loop the target to match was: Lys121-Leu-122-Gly123-Gly124-Ser199. Some notable matches are shown in Table 2:

RMSD % Homology Seq Id. No. Rank Sequence Best fit GLN-VAL-HIS-ASP-GLU 0.154764 68.75 32 81.25 33 2 LYS-GLU-GLY-ASP-LYS 0.15718 9 0.173731 68.75 34 ARG-SER-GLY-ARG-SER THR-LEU-GLY-ASN-SER 0.175554 81.25 35 11 16 0.178772 93.75 36 HIS-PHE-GLY-ALA-GLY

Table 2: Search of V1/V2 loop (Lys121 through Ser199)

15

20

10

5

Based on these searches, constructs encoding Gly-Asn in place of V1/V2 were recommended.

A.3. One Additional Residue Excisions

For a slightly truncated small loop, one more residue was trimmed from each beta strand to slightly shorten the beta sheet. The target to match was: ILE424-ASN425-GLY426-GLY427-LYS432. Results are shown in Table 3:

Table 3: Search of Beta sheet shortened by One residue (Ile424 through Lys432)

. 25

Rank	Sequence	RMSD	% Homology	Seq Id No.
Best fit:	ARG-MET-ALA-PRO-VAL	0.316805	58.33	37
Best	ASP-SER-ASP-GLY-PRO	0.440896	83.33	38
hom:				

Although these searches showed more variation and worse fits than the previous truncation, the Pro-Val or Pro-Leu encoding constructs were very similar. Accordingly, Ala-Pro encoding constructs were recommended.

Sequences encoding gp120 polypeptides having V1/V2 deleted and an additional residue from β -2 or β -3 excised were also searched. The V1/V2 loop the target to match was: VAL120-LYS121-GLY122-GLY123-VAL200. Some notable matches are shown in Table 4.

Table 4: Search of V1/V2 loop (Val120 through Val200)

Rank	Sequence	RMSD	% Homology	Seq Id No
Best fit:	THR-VAL-ASP-PRO-TYR	0.400892	58.33333	39
2	SER-THR-ASN-PRO-LEU	0.402575	54.16667	40
3	THR-ARG-SER-PRO-LEU	0.403965	58.33333	41
7	ARG-MET-ALA-PRO-VAL	0.440118	58.33333	42

15

20

10

5

The construct encoding Ala-Pro (e.g., #7) was recommended.

A.4. Further Excisions

In yet another truncation, an additional residue was trimmed from the β -20 and β -21 strands to further shorten the beta sheet. The target to match was ILE423-ILE424-GLY425-GLY426-ALA433. Notable matches are shown in Table 5.

Table 5: Search of Beta sheet shortened by Two Residues (Ile423 through Ala433)

Rank	Sequence	RMSD	% Homology	Seq Id No
Best fit:	THR-TYR-GLU-GLY-VAL	0.130107	79.16666	43
2	GLN-VAL-GLY-ASN-THR	0.138245	79.16666	44
3:	THR-VAL-GLY-GLY-ILE	0.153362	100	45

25

A construct encoding Gly-Gly (e.g., #3), which has 100% homology, was recommended.

Also searched were sequences encoding a deleted V1/V2 region and at least two residues excised from β -2, β -3 or at least one residue excised from β -2 and β -3. The target to match was: CYS119-VAL120-GLY121-GLY122-ILE201. Notable matches are shown in Table 6.

5

Table 6: Search of V1/V2 loop (Cys119 through Ile201)

Rank	Sequence	RMSD	% Homology	Seq Id No
Best fit:	ASP-LEU-PRO-GLY-CYS	0.250501	75	46
4	ASP-VAL-GLY-GLY-LEU	0.290383	100	47

10

15

It was determined that both constructs would be used.

B.1. Constructs encoding modified Env polypeptides

As described above, the native loops extruding from the 4- β antiparallel-stands were excised and replaced with 1 to 3 residue turns. The loops were replaced so as to leave the entire β -strands or excised by trimming one or more amino acid from each side of the connected strands. The ends of the strands were rejoined with turns that preserve the same backbone geometry (e.g., tertiary structure of β -20 and β -21), as determined by searching the Brookhaven Protein Data Bank.

20

Table 7A is a summary of the truncations of the variable regions 1 and 2 recommended for this study, as determined in Example 1.A. above.

Table 7A

V1/V2 Modifications	SEQ ID NO	Figure
-LEU122-GLY-ASN-SER199	7	10
-LYS121-ALA-PRO-VAL200-	6	9
-VAL120-GLY-GLY-ILE201-	4	7
-VAL120-PRO-GLY-ILE201B-	5	8
-VAL120-GLY-ALA-GLY-ALA204-	3	6
-VAL120-GLY-GLY-ALA-THR202-	8	11
-VAL127-GLY-ALA-GLY-ASN195-	25	28

10

15

5

As previously noted, the polypeptides encoded by the constructs of the present invention are numbered relative to HXB-2, but the particular amino acid residue of the polypeptides encoded by these exemplary constructs is based on SF-162. Thus, for example, although amino acid residue 195 in HXB-2 is a serine (S), constructs encoding polypeptides having then wild type SF162 sequence will have an asparagine (N) at this position. Table 7B shows just three of the variations in amino acid sequence between strains HXB-2 and SF162. The entire sequences, including differences in residue and amino acid number, of HXB-2 and SF162 are shown in the alignment of Figure 2 (SEQ ID NOs:1 and 2).

20

Table 7B

HXB-2 amino acid number	HXB-2 Residue	SF162 Residue/amino acid number
128	Serine (S)	Thr (T)/114
195	Serine (S)	Asn (N)/188
426	Met (M)	Arg (R)/411

25

30

Constructs containing deletions in the β -20 strand, β -21 stand and small loop were also constructed. Shown in Table 8 are constructs encoding truncations in these regions. The constructs in Table 8 are numbered relative to HXB-2 but the unmodified amino acid sequence is based on SF162. Thus, the construct encodes an arginine (Arg) as is found in

SF162 in the amino acid position numbered 426 relative to HXB-2 (See, also, Table 7B). Changes from wildtype (SF162) are shown in bold in Table 8B.

Table 8

Small Loop/β-20 and β-21 (Modified)	SEQ ID NO	Figure
-TRP427-GLY-GLY431-	9	12
-ARG426-GLY-GLY-GLY431-	10	13
-ARG426-GLY-SER-GLY431B-	11	14
-ARG426-GLY-GLY-ASN-LYS432-	12	15
-ASN425-ALA-PRO-LYS432-	13	16
-ILE424-GLY-GLY-ALA433-	14	17
-ILE423-GLY-GLY-MET434-	15	18
GLN422-GLY-GLY-TYR435-	16	19
-GLN422-ALA-PRO-TYR435B-	17	20

15

10

5

The deletion constructs shown in Tables 7 and 8 for each one of the β-strands and combinations of them are constructed. These deletions will be tested in the Env forms gp120, gp140 and gp160 from different HIV strains like subtype B strains (e.g., SF162, US4, SF2), subtype E strains (e.g., CM235) and subtype C strains (e.g., AF110968 or AF110975).

20 Exemplary constructs for SF162 are shown in the

Figures and are summarized in Table 9. As noted above in Figure 2 and Table 7B, in the bridging sheet region, the amino acid sequence of SF162 differs from HXB-2 in that the Met426 of HXB-2 is an Arg in SF162. In Table 9, V1/V2 refers to deletions in the V1/V2 region; # bsm refers to a modification in the bridging sheet small loop.

25

Table 9					
Construct	Seq. Id.	Fig.	Modification/Amino acid sequence		
Val120-Ala204	. 3	6	V1/V2: Val120-Gly-Ala-Gly-Ala204		
Val120-Ile201	4	7	V1/V2: Val120-Gly-Gly-Ile201		
Val120-Ilc201B	5	8	V1/V2: Val120-Pro-Gly-Ile201		
Lys121-Val200	6	9	V1/V2: Lys121-Ala-Pro-Val200		

		T	able 9
Construct	Seq. Id.	Fig.	Modification/Amino acid sequence
Leu122-Ser199	7	10	V1/V2: Leu122-Gly-Asn-Ser199
Vall 20-Thr202	8	11	V1/V2: Val120-Gly-Gly-Ala-Thr202
Trp427-Gly431	9	12	bsm: Trp427-Gly-Gly431
Arg426-Gly431	10	13	bsm: Arg426-Gly-Gly-Gly431
Arg426-Gly431B	11	14	bsm: Arg426-Gly-Ser-Gly431
Arg426-Lys432	12	15	bsm: Arg426-Gly-Gly-Asn-Lys432
Asn425-Lys432	13	16	bsm: Asn425-Ala-Pro-Lys432
Ile424-Ala433	14	17	bsm: Ile424-Gly-Gly-Ala433
Ile423-Met434	15	18	bsm: Ile423-Gly-Gly-Met434
Gln422-Tyr435	16	19	bsm: Gln422-Gly-Gly-Тут435
Val127-Asn195	25	28	bsm: Vall27-Gly-Ala-Gly-Asn195
Gln422-Tyr435B	17	20	bsm: Gln422-Ala-Pro-Tyr435
Leu122-Ser199; Arg426-Gly431	18	21	V1/V2/bsm: Leu122-Gly-Asn-Ser199 Arg426-Gly-Gly-Gly431
Leu122-Ser199; Arg426-Lys432	19	22	V1/V2/bsm: Leu122-Gly-Asn-Ser199 Arg426-Gly-Gly-Asn-Lys432
Leu122-Ser199-Trp427- Gly431	20	23	V1/V2/bsm: Leu122-Gly-Asn-Ser199 Trp427-Gly-Gly431
Lys121-Val200- Asn425-Lys432	21	24	V1/V2/bsm: Lys121-Ala-Pro-Val200 Asn425- Ala-Pro-Lys432
Val120-Ile201-Ile424- Ala433	22	25	V1/V2/bsm: Val120-Gly-Gly-Ile201 Ile424- Gly-Gly-Ala433
Val120-Ile201B-Ile424- Ala433	23	26	V1/V2/bsm: Val120-Pro-Gly-Ile201 Ile424- Gly-Gly-Ala43
Val120-Thr202; Ile424- Ala433	24	27	V1/V2/bsm: Val120-Gly-Gly-Ala-Thr202 Ile424-Gly-Gly-Ala433
Val127-Asn195; Arg426-Gly431	25	29	V1/V2/bsm: Val127-Gly-Ala-Gly-Asn195 Arg426-Gly-Gly-Gly431

Combinations of V1/V2 deletions and bridging sheet small loop modifications in addition to those specifically shown in Table 9 are also within the scope of the present invention. Various forms of the different embodiments of the invention, described herein, may be combined.

The first screening will be done after transient expression in COS-7, RD and/or 293 cells. The proteins that are expressed will be analyzed by immunoblot, ELISA, and for binding to mAbs directed to the CD4 binding site and other important epitopes on gp120 to determine integrity of structure. They will also be tested in a CD4 binding assay and, in addition, the binding of neutralizing antibodies, for example using patient sera or mAb 448D (directed to Glu370 and Tyr384, a region of the CD4 binding groove that is not altered by the deletions).

5

10

The immunogenicity of these novel Env glycoproteins will be tested in rodents and primates. The structures will be administered as DNA vaccines or adjuvanted protein vaccines or in combined modalities. The goal of these vaccinations will be to archive broadly reactive neutralizing antibody responses.

Claims:

What is claimed is:

20

1. A polynucleotide encoding a modified HIV Env polypeptide wherein the polypeptide has at least one amino acid deleted or replaced in the region corresponding to residues 420 to 436 relative to HXB-2 (SEQ ID NO:1).

- 2. The polynucleotide of claim 1, wherein the region corresponding to residues 124-198 relative to HXB-2 is deleted and at least one amino acid is deleted or replaced in the regions corresponding to the residues 119 to 123 and 199 to 210 relative to HXB-2 (SEQ ID NO:1).
- The polynucleotide of claim 1, wherein at least one amino acid in the region
 corresponding to residues 427 through 429 relative to HXB-2 (SEQ ID NO:1) is deleted or replaced.
 - 4. The polynucleotide of claim 2, wherein at least one amino acid of the in the region corresponding to residues 427 through 429 relative to HXB-2 (SEQ ID NO:1) is deleted or replaced.
 - 5. The polynucleotide of claim 1, wherein the amino acid sequence of the modified HIV Env polypeptide is based on strain SF162.
- 25 6. An immunogenic modified HIV Env polypeptide having at least one amino acid deleted or replaced in the region corresponding to residues 420 through 436, relative to HXB-2 (SEQ ID NO:1).
- 7. The polypeptide of claim 6, wherein one amino acid is deleted in the region corresponding to residues 420 through 436, relative to HXB-2 (SEQ ID NO:1).

8. The polypeptide of claim 6, wherein more than one amino acid is deleted in the region corresponding to residues 420 through 436, relative to HXB-2 (SEQ ID NO:1).

- 9. The polypeptide of claim 6, wherein at least one amino acid is replaced in the region corresponding to residues 420 through 436, relative to HXB-2 (SEQ ID NO:1).
 - 10. The polypeptide of claim 6, wherein at least one amino acid residue between about amino acid residue 427 and amino acid residue 429 relative to HXB-2 (SEQ ID NO:1) is deleted or replaced.

11. The polypeptide of claim 6, wherein the V1 and V2 regions of the polypeptide are truncated.

12. The polypeptide of claim 10, wherein the V1 and V2 regions of the polypeptide are truncated.

10

- 13. The polypeptide of claim 6, wherein the amino acid sequence of the modified HIV Env polypeptide is based on strain SF162.
- 20 14. A construct comprising the nucleotide sequence depicted in Figure 6 (SEQ ID NO:3).
 - 15. A construct comprising the nucleotide sequence depicted in Figure 7 (SEQ ID NO:4).
 - A construct comprising the nucleotide sequence depicted in Figure 8 (SEQ ID NO:5).
- 17. A construct comprising the nucleotide sequence depicted in Figure 9 (SEQ IDNO:6).

18. A construct comprising the nucleotide sequence depicted in Figure 10 (SEQ ID NO:7).

- 19. A construct comprising the nucleotide sequence depicted in Figure 11 (SEQ ID5 NO:8).
 - 20. A construct comprising the nucleotide sequence depicted in Figure 12 (SEQ ID NO:9).
- 10 21. A construct comprising the nucleotide sequence depicted in Figure 13 (SEQ ID NO:10).
 - 22. A construct comprising the nucleotide sequence depicted in Figure 14 (SEQ ID NO:11).
 - 23. A construct comprising the nucleotide sequence depicted in Figure 15 (SEQ ID NO:12).
- 24. A construct comprising the nucleotide sequence depicted in Figure 16 (SEQ ID NO:13).

15

- 25. A construct comprising the nucleotide sequence depicted in Figure 17 (SEQ ID NO:14).
- 25 26. A construct comprising the nucleotide sequence depicted in Figure 18 (SEQ ID NO:15).
 - 27. A construct comprising the nucleotide sequence depicted in Figure 19 (SEQ ID NO:16).
 - 28. A construct comprising the nucleotide sequence depicted in Figure 20 (SEQ ID NO:17).

29. A construct comprising the nucleotide sequence depicted in Figure 21 (SEQ ID NO:18).

- 30. A construct comprising the nucleotide sequence depicted in Figure 22 (SEQ IDNO:19).
 - 31. A construct comprising the nucleotide sequence depicted in Figure 23 (SEQ ID NO:20).
- 32. A construct comprising the nucleotide sequence depicted in Figure 24 (SEQ ID NO:21).
 - 33. A construct comprising the nucleotide sequence depicted in Figure 25 (SEQ ID NO:22).
 - 34. A construct comprising the nucleotide sequence depicted in Figure 26 (SEQ ID NO:23).
- 35. A construct comprising the nucleotide sequence depicted in Figure 27 (SEQ ID NO:24).

15

- 36. A construct comprising the nucleotide sequence depicted in Figure 28 (SEQ ID NO:25).
- 25 37. A construct comprising the nucleotide sequence depicted in Figure 29 (SEQ ID NO:26).
 - 38. A vaccine composition comprising a polynucleotide encoding a modified Env polypeptide according to any one of claims 1-5.
 - 39. A vaccine composition comprising a polynucleotide construct encoding a modified Env polypeptide according to any of claims 14-37.

40. A vaccine composition comprising a modified Env polypeptide according to any of claims 6-13.

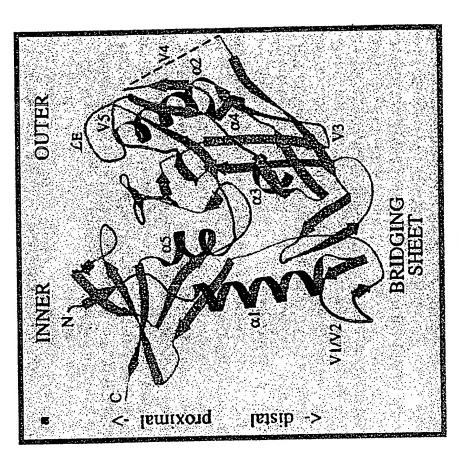
- 41. The vaccine composition of any of claims 38-40, further comprising an adjuvant.
- 42. A method of inducing an immune response in subject comprising, administering a polynucleotide according to any one of claims 1-5 in an amount sufficient to induce an immune response in the subject.

5

- 43. A method of inducing an immune response in subject comprising, administering a polynucleotide construct according to any one of claims 14-37 in an amount sufficient to induce an immune response in the subject.
- 44. A method of inducing an immune response in a subject comprising administering a composition comprising a modified Env polypeptide according to any one of claims 6-13, wherein the composition is administered in an amount sufficient to induce an immune response in the subject
- 45. The method of any of claims 42-44 further comprising administering an adjuvant to the subject.
 - 46. A method of inducing an immune response in a subject comprising
 - (a) administering a first composition comprising a polynucleotide according to any of claims 1-5 in a priming step and
 - (b) administering a second composition comprising a modified Env polypeptide according to any of claims 6-13, as a booster, in an amount sufficient to induce an immune response in the subject.
- 47. The method of claim 46 wherein the first composition or second composition further comprise an adjuvant.

48. The method of claim 46 wherein the first and second compositions further comprise an adjuvant.

gp120 core structure



F/G. 1

WO 00/3	39303		2 /	65	PCT/US99/31272
		1	2 1	03	50
			SupremBt	TENTAL COMPA	
HXB2	(1)	MRVKEKEOHLWRW	PAKAG1 Dr	PROTECT - 2F1 FV	THE STATE OF THE S
162	(1)	MDAMER	Erccarrr	CEANTRSPSEVER	MANAGE AND
SF2	(1)	MKVKGTRRNKOHLWRW	gTLL	LEHLMIC-SETEK	ELARGEMENT (CENTERALIK
CM236	(1)	MRVKETOMN PNLWKW	ĞTL [[Latina IC-Sasnin	ATHROPHY SYLPHOLIC
US4	(1)	MRKHCQHLWRG	GILL	LEELMIC-RETTV	ROMERICA CONTRACTOR K
Consensus	(1)	MRVK YQHLWRW	G TLL	LGMLMIC SATEK	LWVTVYYGVPVWK
		_			
		51 •			100
מענו	(47)		TO THE PARTY OF	west or security of his in in	
HXB2	(47)	EAT REPROASONKAYD	TEAUX	Service assessment of the last	A A CONTRACTOR
162		ERT MAY SASEE KAYD			
SF2		E T EUROPESTIANI Y D			
CM236		Dad Described Colored and the			
US4		EAT			
Consensus	(51)	EATTTLFCASDAKAYD	TEVHNVWA	THACVPTDPNPQE	VVL NVTENFNMW
		101	ال ا		150
HXB2	(97)	(I) DANASON HAD I READ A	elefekk fanswij	SARATA (WASH KOMD)	L
162		ANNAVASYHEDI 63670			
SF2	(96)	Windson Ose I as all	slight (feeting)	SHEEVI NEWD	L
CM236	(96)	ANNERSON OSSESSION			AK
US4		KON COMMON HOLD I SHOW			
Consensus		KNNMVEOMHEDIISLWI			
	,,				
		151		•	200
нхв2	/12E\	KNDTNTNSS	CENTURE.	CET LA ROSE CENT ESTE	
162	(129)	KNATNTKSSI	MANTEND-E	CETVING STATES	IKNEWSKEIALEI
SF2	(134)	GKATNTNESN	MREEFE-R	GUIKNOSENITUS	ROSSINENALER
CM236		LTNVNNITEVS			
US4		GTNSTSGTNETSTNET			
Consensus	(151)	NATNTNSS	KE M K	GEIKNCSFNITTS:	IRDKVQKEYALFY
		201		↓.	• 250
w.no	(170)	· -	TOTAL CENT	Om Program in the Second with	=
HXB2		KGDARNIDNDTTS			
162		KINVV DNDNTS			
SF2		N 13 VV DNAST TNY			
CM236	(179)	KINEVER DNKTSS	SENREINE	AL PATIK CARRAN	XELL TO A CONTRACT OF THE STATE
US4	(191)	K DVV DNDNAS	BRIIN	TENTO AND VE	E STANSING A PAIG
Consensus	(201)	KLDVVPIDND TS	YRLINC	NTSVITQACPKVS	FEPIPIHYCAPAG
		251 ◆ •	•	•	300
HXB2	(223)	F WEEKENN ST RUET 626	TNESTIC	SEPHENI REVEASORS	SSNSSSSSSEVVI
162		FSGUKUNDAK MESER			
SF2		F SOCIETING NO. 17 ENCK SEE			
CM236					
	(220)	COMPACT SELECTION	VANCOUS AND		
US4		FERKCKDEKENETER			
Consensus	(251)	FAILKCNDK FNGTGPO	CTNVSTVQ	CTHGIRPVVSTQL	LLNGSLAEEEVVI
		301	•		350
HXB2	(273)	RSVNFTD (ISKNED) OF	T TO E TOO	THEN WEST KREEK	ORGPGRAFVTISK
162	(266)	RESENTED NEXT PROPERTY OF	KESVE GNG	TENNINERKSTE	GEGRAEYATOD
SF2	(276)	REDNEMNWAREHAROUN	NE SMARNG	RENNARKSLYL	GERALHTTER
CM236		RSENLINNASSEMAVHIN			
US4		RSENFUDNAKENBYOLN			
Consensus		RSENFTDNAKTIIVQLN			
Cousenans	(201)	V2FML I DMWLI I I A ÖFL	TO VELIVE	T TOUNINITADI 1	GEGRAFI IGD

FIG. 2A

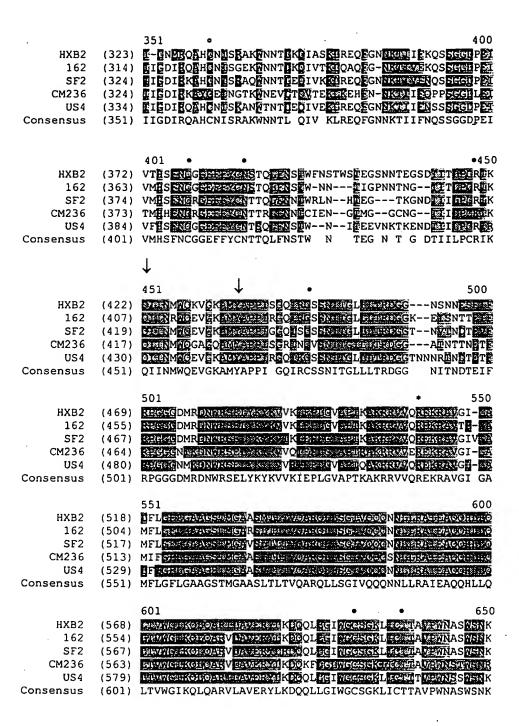


FIG. 2B

FIG. 2C

Consensus (851) AVSLLNATAIAVAEGTDRVIEVAORAFRAILHIPRRIROGLER LL

					DCT/II	S99/31272
WO 00/39303		. 5	/	65	40	377/312/2
		1	* N M C C 1	<u>ATGCAATGAAGAGAGGGCT</u>		
Leu122-Ser199	(1)	GAATTUGULACU	WIGG.	TTGCAATGAAGAGAGGGCT	- TGC 1	
Val127-Asn195	(1)	GAATTCGCCACC	WI GOT	ATGCAATGAAGAGAGGGCT	oreca.	
Val120-Ile201B	(1)	GAATTCGCCACC	WILCO:	ATGCAATGAAGAGAGGGCT(STOCT STOCT	
Vall20-Ala204	(1)	CAATTCCCCACC	vacci Wridii	TIGCAATGAAGAGAGAGGGCT TIGCAATGAAGAGAGAGGCT	TECT	
Val120-Ile201	(1)	CANTILGULACU	MIGG	TGCAATGAAGAGAGGGCT	SACCA STRCI	
Vall20-Thr202		CNATTICCCCACC	ALGG	ŢĠĊAAŢĠĀĀĠĀĠĀĠĠĠĊŢ	TGCT	
Lys121-Va1200	(1)	CANTILL GCCACC	VACC.	ATGCAATGAAGAGAGGGCT	TECT	
Consensus	(1)	41	.AIGG	ATGCAATGAAGAGAGGGCT	. 80	
Leu122-Ser199	(41)	GTGTGCTGCTGC	TGTGT	OCTTTQCTTCTTCGD	CCCAG	
Val127-Asn195	(41)	GTGTGCTGCTGC	TGTGT	GGAGCAGTCTTCGTTTCG	CCAG	
Val120-Ile201B				GGAGCAGTCTTCGTTTCG		
Val120-Ala204	(41)	GTGTGCTGCTGC	TGTGT	GGAGCAGTCTTCGTTTCG	CCAG	
Val120-Ile201		GTGTGCTGCTGC	TGTGT	GGAGCAGTCTTCGTTTCG	CCAG	
Val120-Thr202	(41)	GTGTGCTGCTGC	TGTGT	GGAGCAGTCTTCGTTTCG	CCAG	
Lys121-Val200				GGAGCAGTCTTCGTTTCG		
Consensus	(41)	GTGTGCTGCTGC	TGTGT	GGAGCAGTCTTCGTTTCG	CCAG	
201.521.545	(/	81			120	
Leu122-Ser199	(81)	CGCCGTGGAGAA	GCTGT	GGGTGACCGTGTACTACGC	CGTG	
Val127-Asn195				GGGTGACCGTGTACTACGG		
Val120-Ile201B	(81)	CCCCGTGGAGAA	GCTGT	GGGTGACCGTGTACTACG	CGTG	
Val120-Ala204				GGGTGACCGTGTACTACGC		
Val120-Ile201	(81)			GGGTGACCGTGTACTACG		
Val120-Thr202	(81)			GGGTGACCGTGTACTACGC		
Lys121-Val200		CCCCCTCCAGAA	GCTGT	GGGTGACCGTGTACTACGC	CGTG	
Consensus	(81)	CCCCCTCGAGAA	GCTGT	GGGTGACCGTGTACTACGC	CGTG	
Consensus		121			160	
Leu122-Ser199	(121)	CCCGTGTGGAAG	GAGGC	CÁCCACCACCTGTTCTGC	GCCA	
Val127-Asn195	(121)	CCCGTGTGGAAG	GAGGC	CACCACCACCTGTTCTGC	GCCA	
Val120-Ile201B	(121)	CCCGTGTGGAAG	GAGGC	CACCACCACCCTGTTCTGC	GECA	
Val120-Ala204	(121)	CCCGTGTGGAAG	GAGGC	CACCACCACCTGTTCTGC	GCCA	
Val120-Ile201	(121)	CCCGTGTGGAAG	SAGGO	CACCACCACCTGTTCTGC	GCCA	
Val120-Thr202	(121)	CCCCTCTCCAAG	CAGGO	CACCACCACCCTGTTCTGC	GCCA	
Lvs121-Val200	(121)	CCCGTGTGGAAG	SAGGÇ	CACCACCACCTGTTCTGC	GCCA	
Consensus	(121)	CCCGTGTGGAAG	GAGGC	CACCACCACCTGTTCTGC	GCCA	
		161			200	
Leu122-Ser199	(161)	GCGACGCCAAGG	CTAC	GACACCGAGGTGCACAACC	TGTG	
Val127-Asn195	(161)	GCGACGGGAAGG	CTAC	GACACCGAGGTGCACAACC	TGTG	
Val120-Ile201B	(161)	GCGACGCCAAGG	CTAC	GACACCGAGGTGCACAACG	TGTG	
Val120-Ala204	(161)	GCGACGCCAAGGC	CTAC	GACACCGAGGTGCACAACC	TGTG	
Val120-Ile201	(161)	GCGACGCCAAGGC	CTAC	GAÇACCGAGGTGCACAACC	TGTG	
Val120-Thr202	(161)	GCGACGCCAAGGC	CTAC	GACACCGAGGTGCACAACG	TGTG	
Lys121-Val200	(161)			GACACCGAGGTGCACAACG		
Consensus	(161)			GACACCGAGGTGCACAACG		
•••••		201			240	
Leu122-Ser199	(201)	GGCCACCCACGCC	TGCG	TGECCACCGACCCCAACCC	CCAG	•
Val127-Asn195	(201)	GGCCACCCACGCC	TGCG	TGCCCACCGACCCCAACCC	CCAG	
Val120-Ile201B	(201)	GCCCACCCACGCC	TGCG	TGCCCACCGACCCCAACCC	CCAG	
Val120-Ala204	(201)	GGCCACCCACGCC	TGČG	TGCCCACCGACCCCAACCC	CCAG	
Val120-Ile201	(201)	GGCCACCCACGCC	TGCG	TGCCCACCGACCCCAACCC	CCAG	
Val120-Thr202	(201)	GGCCACCCACGCC	TGCG	TGCCCACCGACCCCAACCC	CCAG	
Lys121-Val200	(201)	GGCCACCCACGCC	TGCG	TGCCCACCGACCCCAACCC	CCAG	
Consensus	(201)	GGCCACCCACGCC	TGCG	TGCCCACCGACCCCAACCC	CCAG	
00.10011000	/	241			280	
Leul22-Ser199	(241)		AGAA	CGTGACCGAGAACTTCAAC		
Val127-Asn195	(241)	GAGATCGTGCTGC	AGAA	CGTGACCGAGAACTTCAAC	ATGT	
.dil. noniyo	(= 3 1)	Saladit sacreta ap	:	ਲਾਮ ਦਾ ਜਾਣ ਨਾਲ ਹਨਾਂ ਜਾਣ ਦੇ ਦੇ ਜੀ ਜੈ ਜਿ		

(439) AAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCCTGCA

Val120-Ile201

```
(439) AAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCCTGCA
 Val120-Thr202
                  (445) AAGTGCAACGACAAGAAGTTCAACGGCAGCGCCCCTGCA
 Lys121-Val200
                  (481) AAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCCTGCA
     Consensus
                  (491) CCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCCC
 Leu122-Ser199
                  (521) CCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCCC
 Val127-Asn195
                  (479) CCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCCC
Val120-Ile201E
                  (473) CCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCCC
 Val120-Ala204
                  (479) CCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCCC
 Val120-Ile201
                  (479) CCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCCC
 Val120-Thr202
                  (485) CCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCCC
 Lys121-Val200
                  (521) CCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCCC
     Consensus
                  (531) CGTGGTGAGCACCCAGCTGCTGAACGGCAGCCTGGCC
 Leu122-Ser199
                  (561) CGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCC
 Val127-Asn195
                  (519) CGTGGTGAGCACCCAGCTGCTGGTGAACGGCAGCCTGGCC
(513) CGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCC
Val120-Ile201B
 Val120-Ala204
                  (519) CGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCC
 Val120-Ile201
                  (519) CGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCC
 Val120-Thr202
                  (525) CGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCC
 Lys121-Val200
                  (561) CGTGGTGAGCACCCAGCTGCTGCAACGGCAGCCTGGCC
     Consensus
                  (571) GAGGAGGGCGTGGTGATCCGCAGCGAGAACTTCACCGACA
 Leu122-Ser199
                  (601) GAGGAGGGCGTGGTGATCCGCAGCGAGAACTTCACCGACA
 Val127-Asn195
                        GAGGAGGCGTGGTGATCCGCAGCGAGAACTTCACCGACA
Val120-Ile201B
                  (553) GAGGAGGGCGTGGTGATCCGCAGCGAGAACTTCACCGACA
 Val120-Ala204
                  (559) GAGGAGGGCGTGGTGATCCGCAGCGAGAACTTCACCGACA
 Val120-Ile201
                  (559) GAGGAGGCGTGGTGATCCGCAGCGAGAACTTCACCGACA
 Val120-Thr202
                  (565) GAGGAGGGCGTGGTGATCCGCAGCGAGAACTTCACCGACA
Lys121-Val200
                  (601) GAGGAGGCGTGGTGATCCGCAGCGAGAACTTCACCGACA
     Consensus
                  (611) ACGCCAAGACCATCATCGTGGAGCTGAAGGAGAGCGTGGA
 Leu122-Ser199
                        ACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGA
                  (641)
 Val127-Asn195
                  (599) ACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGA
(593) ACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGA
Val120-Ile201B
 Val120-Ala204
                  (599) ACCCCAAGACCATCATCGTGCAGGTGAAGGAGAGCGTGGA
 Val120-Ile201
                  (599) ACGCCAAGACCATCATEGTGCAGCTGAAGGAGCGTGGA
 Val120-Thr202
                  (605) ACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGA
 Lys121-Val200
                  (641) ACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGA
     Consensus
                  (651) GATCAACTGCACCCGCCCAACAACAACACCCGCAAGAGC
 Leul22-Ser199
                 (681) GATCAACTGCACCCGCCCAACAACAACACCCCGCAAGAGC
 Val127-Asn195
                 (639) GATCAACTGCACCCGCCCAACAACAACACCCCGCAAGAGC
Val120-Ile201B
                 (633) GATCAACTGCACCCGCCCAACAACAACACCCGCAAGAGC
 Val120-Ala204
                  (639) GATCAACTGCACCCGCCCCAACAACAACACCCGCAAGAGC
 Val120-11e201
                 (639) GATCANCTECACCEGCCCGAACAACAACACCCGCAAGAGC
(645) GATCANCTGCACCCGCCCCAACAACAACACCCGCAAGAGC
 Val120-Thr202
Lvs121-Val200
                  (681) GATCAACTGCACCCGCCCAACAACAACACCCCGCAAGAGC
     Consensus
                 (691) ATCACCATCGGCCCGGCCGCGCCTTCTACGCCACCGGCG
Leu122-Ser199
                 (721) ATCACCATCGGCCCGGCCGCCCTTCTACGCCACCGGCG
Val127-Asn195
                  (679) ATCACCATCGGCCCCGGCCGCCTTCTACGCCACCGGCG
Val120-Ile201B
                 (673) ATCACCATCGGCCCGGCCGCCCTTCTACGCCACCGGCG
 Val120-Ala204
                 (679) ATCACCATCGGCCCGGCCGCCCTTCTACGCCACCGGCG
Val120-Ile201
                  (679) ATCACCATCGGCCCCGGCCGCCTTCTACGCCACCGGCG
 Val120-Thr202
                  (685) ATCACCATCGGCCCCGGCCGCCCTTCTACGCCACCGGCG
Lys121-Val200
                 (721) ATCACCATCGGCCCGGCCGCCTTCTACGCCACCGGCG
     Consensus
```

```
Val120-Ala204
                 (959) ACAACACCAACGGCACCATCACCCTGCCCTGCCGCATCAA
 Val120-Ile201
                 (959) ACAACACCAACGGCACCATCACCCTGCCCTGCCGCATCAA
 Val120-Thr202
                 (965) ACAACACCAACGGCACCATCACCCTGCCCTGCCGCATCAA
 Lys121-Val200
                (1001) ACAACACCAACGGCACCATCACCCTGCCCTGCCGCATCAA
     Consensus
                (1011) GCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATG
 Leu122-Ser199
                (1041) GCAGATCATCAACCGCTGGCAGGTGGGCAAGGCCATG
 Val127-Asn195
                 (999) GCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATG.
Val120-Ile201B
                 (993) GCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATG
 Val120-Ala204
                 (999) GCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATG
 Val120-Ile201
                 (999) GCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATG
Val120-Thr202
                (1005) GCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATG
Lys121-Val200
                (1041) GCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATG
     Consensus
                       1081
                (1051) TACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCA
Leu122-Ser199
                (1081) TACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCA
 Val127-Asn195
                (1039) TACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCA
Val120-Ile201B
                (1033) TACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCA
 Val120-Ala204
                (1039) TACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCA
Val120-Ile201
                      TACGCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCA
Val120-Thr202
                (1039)
                (1045) TACGECCCCCCATCCGCGCCAGATCCGCTGCAGCAGCA
Lys121-Val200
                (1081) TACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCA
     Consensus
                (1091) ACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAAGGA
Leu122-Ser199
                (1121) ACATCACCGGCCTGCTGCTGACCCGCGACGGCGACGGCAAGGA
Val127-Asn195
                (1079) ACATCACCGGCCTGCTGACCCGCGACGGCGCAAGGA
Val120-Ile201B
                (1073) ACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAAGGA
Val120-Ala204
                (1079) ACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAAGGA
Val120-Ile201
                (1079) ACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAAGGA
Val120-Thr202
               (1085) ACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAAGGA
Lys121-Val200
               (1121) ACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAAGGA
               (1131) GATCAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGGC
Leul22-Ser199
               (1161) GATCAGCAACACCACCGAGATCTTCCGCCCCGGGGGGGGC
Val127-Asn195
                      Val120-Ile201B
                (1119)
                (1113) GATCAGCAACACGACGAGATCTTCCGCCCGGCGGCGGC
Val120-Ala204
                (1119) GATCAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGGC
Val120-Ile201
               (1119) GATCAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGCC
Val120-Thr202
               (1125) GATCAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGC
Lys121-Val200
               (1161) GATCAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGGC
    Consensus
               (1171) GACATGCGCGACAACTGGCGCAGCGAGCTGTACAAGTACA
Leul22-Ser199
                      GACATGCGCGACAACTGGCGCGCGGGCTGTACAAGTACA
Val127-Asn195
               (1201)
               (1159) GACATGCGCGACAACTGGCGCAGCGAGCTGTACAAGTACA
Val120-Ile201B
               (1153) GACATGCGCGACAACTGGCGCAGCGAGCTGTACAAGTACA
Val120-Ala204
               (1159) GACATGCGCGACAACTGGCGCGAGCTGTACAAGTACA
Val120-Ile201
               (1159) GACATGCGCGACAACTGGCGCAGCGAGCTGTACAAGTACA
Val120-Thr202
               (1165) GACATGCGCGACAACTGGCGCAGCGAGCTGTACAAGTACA
Lys121-Va1200
               (1201) GACATGCGCGACAACTGGCGCAGCGAGCTGTACAAGTACA
    Consensus
               (1211) AGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCACCAA
Leu122-Ser199
               (1241) AGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCACCAA
Val127-Asn195
               (1199) AGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCACCAA
Val120-Ile201B
               (1193) AGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCACCAA
Val120-Ala204
               (1199) AGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCACCAA
Val120-Ile201
```

```
1521
                 (1491) GCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTG
 Leu122-Ser199
                 (1521) GCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTG
 Val127-Asn195
                 (1479) GCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTG
Val120-Ile201B
                 (1473) GCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTG
 Val120-Ala204
                 (1479) GCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTG
 Val120-Ile201
                 (1479) GCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTG
 Val120-Thr202
                 (1485) GCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTG
 Lys121-Val200
     Consensus
                 (1521) GCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTG
                 (1531) GGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCG
 Leu122-Ser199
                 (1561) GGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCG
(1519) GGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCG
 Val127-Asn195
Vall20-Ile201B
                 (1513) GGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCG
 Val120-Ala204
                 (1519) GGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCG
 Val120-Ile201
                 (1519) GGCATCTGGGGCTGCAGCGGCAAGGTGATCTGCACCACCG
 Val120-Thr202
                 (1525) GGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCG
 Lys121-Val200
                 (1561) GGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCG
     Consensus
                 (1571) CCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGA
 Leu122-Ser199
                 (1601) CCGTGCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGA
 Val127-Asn195
                 (1559) CCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGA
(1553) CCGTGCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGA
Val120-Ile201B
 Val120-Ala204
                 (1559) CCGTGCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGA
 Val120-Ile201
                 (1559) CCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGA
 Val120-Thr202
                 (1565) CCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGA
 Lys121-Val200
                 (1601) CCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGA
     Consensus
                         1641
                 (1611) CCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGC
 Leu122-Ser199
                 (1641) CCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGC
 Val127-Asn195
                 (1599) CCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGC
Val120-Ile201B
                 (1593) CCAGATCTGGAAGAACATGACCTGGATGGAGTGGGAGCGC
 Val120-Ala204
                 (1599) CCAGATCTGGAAGAACATGACCTGGATGGAGTGGGAGCGC
 Val120-Ile201
                 (1599) CCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGC
 Val120-Thr202
                (1605) CCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGC
 Lys121-Val200
                (1641) CCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGC
     Consensus
                 (1651) GAGATCGACAACTACACCAACCTGATCTACACCCTGATCG
(1681) GAGATCGACAACTACACCAACCTGATCTACACCCTGATCG
 Leu122-Ser199
 Val127-Asn195
                (1639) GAGATCGACAACTACACCAACCTGATCTACACCCTGATCG
Val120-Ile201B
                 (1633) GAGATCGACAACTACACCAACCTGATCTACACCCTGATCG
 Val120-Ala204
                 (1639) GAGATEGACAACTACACCAACCTGATCTACACCCTGATCG
 Val120-Ile201
                 (1639) GAGATEGACAACTACACCAAGCTEATETACACCCTGATCG
 Val120-Thr202
                 (1645) GAGATCGACAACTACACCAACCTGATCTACACCCTGATCG
Lys121-Val200
                 (1681) GAGATCGACAACTACACCAACCTGATCTACACCCTGATCG
     Consensus
                (1691) AGGAGAGCCAGAACCAGCAGGAGAACGAGCAGGAGCT
 Leu122-Ser199
                (1721) AGGAGAGCCAGAACCAGCAGGAGAACGAGCAGGAGCT
 Val127-Asn195
               (1679) AGGAGAGCCAGAACCAGCAGGAGAACGAGCAGGAGCT
Val120-Ile201B
                (1673) AGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCT
 Val120-Ala204
Val120-Ile201
                (1679) AGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCT
                (1679) AGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCT
Val120-Thr202
                (1685) AGGAGAGCCAGAACCAGCAGGAGAACGAGCAGCAGGAGCT
Lys121-Val200
                (1721) AGGAGAGCCAGAACCAGCAGGAGAACGAGCAGGAGCT
     Consensus
                        1761
                (1731) GCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTC
Leu122-Ser199
                (1761) GCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTC
```

```
Vall20-Ile201B (1719) GCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTC
                 (1713) GCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTC
  Val120-Ala204
                 (1719) GCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTC
  Val120-Ile201
  Val120-Thr202
                 (1719) GCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTC
  Lys121-Val200
                 (1725) GCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTC
                 (1761) GCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTC
      Consensus
                 (1771) GACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCA
  Leu122-Ser199
  Val127-Asn195
                 (1801) GACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCA
                 (1759) GACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCA
(1753) GACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCA
 Val120-Ile201B
  Val120-Ala204
  Val120-Ile201
                 (1759) GACATCAGCAAGTGGCTGTGGTAGATCAAGATCTTCATCA
 Val120-Thr202
                 (1759) GACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCA
  Lys121-Va1200
                (1765) GACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCA
     Consensus
                (1801) GACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCA
                        1841
 Leul22-Ser199
                (1811) TGATCGTGGGGGGCCTGGTGGGCCTGCGCATCGTGTTCAC
                (1841) TGATCGTGGGCGGCCTGGTGGGCCTGCGCATCGTGTTCAC
 Val127-Asn195
Val120-Ile201B
                (1799) TGATCGTGGGCGGCCTGGTGGGCCTGCGCATCGTGTTCAC
 Val120-Ala204
                (1793) TGATCGTGGGCGGCCTGGTGGGCCTGCGCATCGTGTTCAC
                 (1799) TGATCGTGGGCGGCCTGGTGGGCCTGCGCATCGTGTTCAC
 Val120-Ile201
 Val120-Thr202
                 (1799) TGATCGTGGGCGGCCTGGTGGGCCTGCGCATCGTGTTCAC
 Lys121-Va1200
                 (1805) TGATCGTGGGGGGCCTGGTGGGCCTGCGCATCGTGTTCAC
     Consensus
                (1841) TGATCGTGGGCGGCCTGGTGGGCCTGCGCATCGTGTTCAC
                        1881
 Leu122-Ser199
                (1851) CGTSGTGAGCATCGTGAACCGGGTGCGCCAGGGCTACAGC
 Val127-Asn195
                (1881) CGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGC
               (1839) CGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGC
Val120-Ile201B
 Val120-Ala204
                (1833) CGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGC
                (1839) CGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGC
 Val120-Ile201
 Val120-Thr202
                (1839) CGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGC
 Lys121-Va1200
                (1845) CGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGC
     Consensus
                (1881) CGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGC
                       1921
                                                            1960
 Leu122-Ser199
                (1891) CCCCTGAGCTTCCAGACCCGCTTCCCCGCGCCCCCCGCGCCC
 Val127-Asn195
                (1921) GCCCCGAGGCTTCGAGACCGGCTTTCCCGGCCCCCCGCGGCCC
Val120-Ile201B
                (1879) CECETGACETTCCACACEGGGTTEECEGGCCCCGGGGCC (1873) CECETGAGCTTCCAGACCGGTTEECEGGCCCCCGCGGCC
 Val120-Ala204
 Val120-Ile201
                (1879) CCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCCCGCGGCC
 Val120-Thr202
                (1879) CCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCCGCGCCCC
 Lys121-Val200
               (1885) CCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCCGCGGCC
     Consensus
                (1921) CCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCCGCGGCC
               (1931) CCGACCGCCCGAGGGCATCGAGGAGGAGGGCGGCGAGCG
 Leu122-Ser199
 Val127-Asn195
Val120-Ile201B
                (1919)
                       (1913) CCGACGGCGCGAGGGCATCGAGGAGGAGGGCGCGAGCG
Val120-Ala204
                Val120-Ile201
                (1919) CCGACCGCCCCGAGGGCATCGAGGAGGAGGGCGCGAGCG
 Val120-Thr202
Lys121-Val200
                (1925) CCGACCGCCCGAGGGCATCGAGGAGGAGGGCGCGAGCG
                (1961) CCGACCGCCCGAGGGCATCGAGGAGGAGGGCGGCGAGCG
     Consensus
Leul22-Ser199
               (1971) CGACCGCGACCGCAGCAGCCCCTGGTGCACGGCCTGCTG
Val127-Asn195
               (2001) CGACCGCGACCGCAGCAGCCCCTGGTGCACGGCCTGCTG
Val120-Ile201B
               (1959) CGACCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTG
Val120-Ala204
               (1953) CGACCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTG
Val120-Ile201 (1959) CGACCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTG
```

PCT/US99/31272 WO 00/39303

```
Vall20-Thr202 (1959) CGACCGCGACCGCAGCGCCCCTGGTGCACGGCCTGCTG
                  (1965) CGACCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTG
  Lys121-Val200
      Consensus
                  (2001) CGACCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTG
                          2041
                 (2011) GCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
  Leu122-Ser199
                 (2041) GCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
  Val127-Asn195
 Vall20-Ile201B (1999) GCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
                  (1993) GCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
  Val120-Ala204
  Val120-Ile201
                  (1999) GCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
  Val120-Thr202
                  (1999) GCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
                  (2005) GCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
  Lys121-Val200
      Consensus (2041) GCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
  Leu122-Ser199
                 (2051) GCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCCG
                 (2081) GCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCCG
  Val127-Asn195
                 (2039) GCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCGG
(2033) GCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCCGG
(2039) GCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCCGG
Val120-Ile201B
  Val120-Ala204
 Val120-Ile201
                  (2039) GCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCCG
 Val120-Thr202
 Lvs121-Va1200
                 (2045) GCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCCG
      Consensus
                 (2081) GCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCCG
                          2121
                 (2091) CATCGTGGAGCTGCTGGGCCGCGGGCTGGGAGGCCCTG
 Leu122-Ser199
                  (2121) CATCGTGGAGCTGCTGGGCCGCGGCTGGGAGGCCCTG
 Val127-Asn195
                 (2079) CATGETGGAGETGCTGGGGCGGCGGGGCTGGGAGGCCCTG
(2073) CATGGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGCCCTG
Val120-Ile201B
 Val120-Ala204
                 (2079) CATCGTGGAGCTGCAGGGCCGCCGGCTGGGAGGCCCTG
 Val120-Ile201
                 (2079) CATCGTGGAGCTGCTGGGCCGCGGGCTGGGAGGCCCTG
 Val120-Thr202
                 (2085) CATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGGCCCTG
 Lys121-Val200
                 (2121) CATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGGCCCTG
      Consensus
                (2131) AAGTAGTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGC
(2161) AAGTACTGGGGGCAAGCTGCTGCAGTACTGGATCCAGGAGC
 Leu122-Ser199
 Val127-Asn195
                (2119) AAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGC
Val120-Ile201B
                 (2113) AAGUAGIEGGGAAAGSTEGUIGGACTACUGGATECCAGGAGC
 Val120-Ala204
                 (2119) AAGTACTEGEGCAAGCTGCTGCAGTACTGGATCCAGGAGC
(2119) AAGTACTGGGGCAAGCTGCTGCAGTACTGGATCCAGGAGC
 Val120-Ile201
 Val120-Thr202
                 (2125) AAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGC
 Lys121-Val200
     Consensus
                 (2161) AAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGC
                         2201
                                                                 2240
 Leu122-Ser199
                (2171)
                         TGAAGAAGAGGGCGTGAGCCTGTTGGAGGCGATCGCCAT
                 (2201) TGATGATCAGEGCEGTGAGECTGTTCGAGGCCATCGCCAT
(2159) TGATGATCAGEGCTGTGAGCCTGTTCGACGCCATCGCCAT
 Val127-Asn195
Val120-Ile201B
                 (2153) TGAAGAACAGEGCCETTGAGCCTG ENEGACGCCATCGCCAT
 Val120-Ala204
 Val120-Ile201
                 (2159) TGAAGAACAGEGECCTGAGCCTGTTCGACGCCATCGCCAT
                        TGAAGAACAGCGCGTGAGCGTGTTCGACGCCATCGCCAT
 Val120-Thr202
                 (2159)
 Lys121-Val200
                 (2165) TGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCAT
                (2201) TGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCAT
     Consensus
                         2241
                (2211) CGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCC
 Leul22-Ser199
 Val127-Asn195 (2241) CGCCGTGGCCGAGGGCACCGCATCATCGAGGTGGCC
Val120-Ile201B (2199) CGCCGTGGCCGAGGGCACCGCATCATCGAGGTGGCC
                (2193) CGCCGTGGCCGAGGGCACCGCATCATCGAGGTGGCC
 Val120-Ala204
                 (2199) CGCCGTGGCCGAGGGCACCGCATCATCGAGGTGGCC
 Val120-Ile201
 Val120-Thr202
                 (2199) CGCCGTGGCCGAGGGCACCGCATCATCGAGGTGGCC
                (2205) CGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCC
Lys121-Val200
     Consensus (2241) CGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCC
```

```
2281
                  (2251) CAGCGCATCGGCCGCGCCTTCCTGCACATCCCCCGCCGCA
 Leu122-Ser199
                  (2281) CAGCGCATCGGCCGCGCCTTCCTGGACATCCCCCGCCGCA
 Val127-Asn195
                  (2239) CAGEGCATCGGEGGCGCCTTCCTGCACATCCCCCGCCGCA
Val120-Ile201B
                  (2233) CAGCGCATCGGCCGCCCTTCCTGCACATCCCCCGCCGCA
 Val120-Ala204
                  (2239) CAGCCCATCGGCCGCCCTTCCTGCACATCCCCCGCCGCA
 Val120-Ile201
                  (2239) CAGCGCATCGGCCGCGCCTTCCTGCACATCCCCCGCCGCA
 Val120-Thr202
                  (2245) CAGCGCATCGGCCGCGCCTTCCTGCACATCCCCCGCCGCA
 Lys121-Val200
                  (2281) CAGCGCATCGGCCGCGCCTTCCTGCACATCCCCCGCCGCA
     Consensus
                         2321
                                                                 2360
                  (2291) TCCCCCAGGGCTTCGAGCGCGCCCTGCTGAACTCGAGCG
 Leu122-Ser199
                  (2321) TCCGCCAGGGCTTCGAGCGCGCCCTCCTGTAACTCGAG--
 Val127-Asn195
                 (2279) TCCGCCAGGCTTCGAGCGCCCTGCTGTAACTCGAGCG
(2273) TCCGCCAGGCCTTCGAGGGCGCGCGCTGCTGTAACTCGAG--
(2279) TCCGCCAGGGCTTCGAGCGCGCGCGCTGCTAACTCGAG--
Val120-Ile201B
 Val120-Ala204
 Val120-Ile201
                 (2279) TCCGCCAGGGCTTCGAGCGCGCCCTGCTGAACTCGAG--
 Val120-Thr202
                 (2285) TCCGCCAGGGCTTCGAGCGCCCCTGCTGTAACTCGAGCG
 Lys121-Val200
                 (2321) TCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACTCGAG
     Consensus
                         2361
                 (2331) TGCT
 Leu122-Ser199
                 (2359) ----
 Val127-Asn195
Vall20-Ile201B
                 (2319) TGCT
 Val120-Ala204
                 (2311) ----
                 (2317) ----
 Val120-Ile201
 Val120-Thr202
                 (2317) ----
                 (2325) TGCT
 Lys121-Val200
                 (2361)
     Consensus
```

WO 00/39303		15	1	65	PCT/US99/31272 40
Ile424-Ala433	(1)	opposite de la constant	त्य:\स्ट्रेल्डाः\	Refeliations	<u>्रित्रीत्तर्भतित्रोत्रकारेत्यत्रो</u>
Trp427-Gly431					CARCACIONICACION
Gln422-Tyr435B	(1)	ल १५११मे हेल्ल्ला व्यक्त	ભુગ્ય લ્લા એ	elictivitatetiv	(त्रे:\ल्यात्त्वात्त्व्यं स्टा <u>प्ट</u> ार्
Arg426-Gly431					ोलगलग्रहासंस्थापस्य संस्था
Ile423-Met434					<u>rengestred els reage Appearan</u>
Gln422-Tyr435	(1)				<u>किनेक्सेट्राव्यक्षेत्रहत्त्रियः स</u>
Arg426-Lys432	(1)	संक्रिक्सेसेस्वर्वर स्था	בייוי ויבי ביי	AUGUSTINES .	/ca:/@j:/cicletthicgifethic
Arg426-Gly431B	(1)	विकार कार्य स्ट्रेसिक होते	e vilejci.V	لأداخ عربها والأحدية	(देशक्षेत्रदेशक्षेत्रकृतिकार्
Asn425-Lys432	(1)				(a) a race on the race
Consensus	(1)	GAATTCGCCACO	CATGGAT	GCAATGA!	GAGAGGGCTCTGCT 80
Ile424-Ala433	(41)	Chi cita felantel	देश देश देश देश हैं देश देश हैं के देश हैं कि कि का कि	त्तुः)दाकः (सः हः	नस्र हेल्ल्स्कृत्सार् लेखल्ल्स्सर्ट
Trp427-Gly431	(41)				म्बर्ग्यद्वारम्बद्धाः स्टब्स्य स्टब्स्य स्टब्स्य स्टब्स्य स्टब्स्य स्टब्स्य स्टब्स्य स्टब्स्य स्टब्स्य स्टब्स्य स्टब्स्य स्टब्स्य स
Gln422-Tyr435B	(41)	CHICHTE MERCHE	Sept Child	d:/de/e	मासद्रिक्त दमासीमद्र ना ल्ल इत्तर हो
Arg426-Gly431	(41)				HARACTER COLLEGE
Ile423-Met434	(41)				THE CARRESPECTIONS TO
Gln422-Tyr435	(41)				ning sieh menteralakaistrie
Arg426-Lys432	(41)	देशहरकश्चीत्रहित्रकृतिक	alter et et et	terricles relation	:::::::::::::::::::::::::::::::::::::
Arg426-Gly431B	(41)				(हेन्द्र ाट्स शामाने सहाराष्ट्र १९
Asn425-Lys432	(41)				असाक्षतामा ११ तालक व ंदाह
Consensus	(41)	GTGTGCTGCTGC	CTGTGTG	GAGCAGTO	TTCGTTTCGCCCAG 120
Ile424-Ala433	(81)	्रह्में देश होता है लि र्युक्त है।	(GL) (C) (C	lefon Ginger	<u>ः स्वाप्तरमान्त्राचान्द्रवरम्</u>
Trp427-Gly431	(81)				।इंदेहरीया क्षेत्र के क्षेत्र के क्षेत्र के क्षेत्र के क्षेत्र के कि
Gln422-Tyr435B	(81)				त्रिक्षकृतस्य स्टब्स्ट्रिल्ला
Arg426-Gly431	(81)	लेसिट लंदनेत्रधास्त्रकार्या	Security is	<i>ल्ल्या के स</i> र्वे	त्रद्धाः अत्यक्षायः स्टब्स् स्टब्स् इत्याच्या
Ile423-Met434	(81)				१८८ को १८४८ (स्टे स्ट्रेस्ट्रेस्ट्रेस्ट्र
Gln422-Tyr435	(81)				संदर्भगटबस्यान्। संदर्भट स्ट
Arg426-Lys432					ecsestories elected
Arg426-Gly431B	(81)				ं हो। इ.स. इ.स. इ.स. हे ने इ.स. हे के किए हैं कि किए हैं कि किए ह
Asn425-Lys432			(Gentering)	<u>इस्ट्रिक्ट इस्ट्रोस्ट</u>	ĸĠĸĸĸĸĸĸĸĸ
Consensus	(81)	CGCCGTGGAGAA	GCTGTG	GGTGACCG	TGTACTACGGCGTG 160
· Ile424-Ala433	(121)	accelenses/cicles se	144.144.40	otolerite este	राक्ष्मिस्सारक संस् लिय क्ष ा
Trp427-Gly431	(121)	<i>सम्मृहदर्शसम्बद्धाः</i>	(e):\c\r\r\r\c	ing control and	CETTATES POSTCOCK
Gln422-Tyr435B	(121)				प्रदेश हैं का श्रेम (द्वेष द ्वार के
Arg426-Gly431	(121)	व्यवस्थान विकास	(E) (E) C C C	preferintestas	्रामास्यक्षकर्वन्यस्य स्थान
Ile423-Met434	(121)	लंबास कार्य स्था संबंधित है।	()(c) e/c; s	15(0) (15) (82.5)	ેલ્ફા'લ મેં જ્યારિકાલ એ <u>.</u>
Gln422-Tyr435	(121)	अंदर्ग सम्मन्द्रा स्ट्रास	(d+(d;d;))	ofe chile exile	्रहार्थक क्षेत्र के प्रश् <u>य दिल</u> ्ला
Arg426-Lys432	(121)	eletelete careletario	(Stitelettion	etologitolegito	લુડનું મહાન મામ કરવાનું લાલું છે. જો
Arg426-Gly431B	(121)	waterand-specification	MAGE STO	the editional po	हात्र द्वार इंड्यम् व हेष्ट्र हो।
Asn425-Lys432	(121)	distribution (clos). To	(कर्म्स्ट्रॉल्ड्रॉट)	the circonie	सन्दर्भ संदर्भ का स्वासन्दर्भ करें के अपने का
Consensus		161			CCTGTTCTGCGCCA 200
Ile424-Ala433					quiqes har tracencence
Trp427-Gly431					enterentent den graft
Gln422-Tyr435B					elike 19:koza vojenka vje
Arg426-Gly431					enneles feith to enne.
Ile423-Met434					इ.स.च.स.स.स.स.स.स.स. .
Gln422-Tyr435					enkcięł i czał odcił i c Kę
Arg426-Lys432					enterological desirente
Arg426-Gly431B	(161)				dentification of the state of t
Asn425-Lys432	(161)				के त्रोलं श ्चिम्
Consensus	(161)	GCGACGCCAAGG 201	CCTACG	ACACCGAG	GTGCACAACGTGTG 240
Ile424-Ala433	(201)	बंबरामग्रह्मकोग्रिवे	्रोत्रद्धाः द्रम	स ् रिल्लाकाटी	radicity fordescrie

FIG. 4A

FIG. 4B

(401) रिवदायकाम्बद्धराज्यात्वार्भेकार्यकार्भेकार्यकार्भवार्भेवार्भेवार्भेवार्भेवार्भेवार्भे

Gln422-Tyr435B

WO 00/39303		17	/ 65	PCT	US99/31272
Arg426-Gly431	(401)	Activitation and		je pretet strejet y ett pike je ji	
Ile423-Met434	(401)			remediation remedia	
Gln422-Tyr435	(401)			/कोदेव्यक्तिकातिकातिकातिकातिका	
Arg426-Lys432	(401)			रद्धार स्थापना स्थापन स्थाप स्थापन स्थापन	
Arg426-Gly431B	(401)			leineen vieraleranden.	
Asn425-Lys432	(401)				
Consensus	(401)			Check recognition and the control of	
	(401)	441	ACCAAGAGCAGCAA	CTGGAAGGAGATGGA 480	
Ile424-Ala433	(441)		hotalotalicaniciosilele	भूभू <mark>लक्ष्</mark> रीयद्वारतागुरुद्वगुरुद्	
Trp427-Gly431	(441)	राहित्रहात्वरास्त्रां स्थाप	तक्षा संदर्भाय संदर्भात	माम्बर्भारत स्वयवस्थल -	-
Gln422-Tyr435B	(441)	ब्रह्मेद्रोदेव्द्रहे व्यक्ष्यं क्ष्यं	(क्रे:भ्रोदोश्याक्रेस्ट्रेल्येस्ट्र	apple properties and the particular and the particu	
Arg426-Gly431	(441)	लेखेलालच्छालाले हेन्स्	nestation strong cities	printeration desires and	
Ile423-Met434	(441)	्ट्रिक्स संस्थित वस्ति होत	leastle stickle det	makely the saccast of care co	
Gln422-Tyr435	(441)		antigations and	रामसंस्कृतेलाक्षरकारे ला स्ट्राल्ड	
Arg426-Lys432	(441)	elojete eleleletitetik	Contraction to the content	शक्ताहरू हो द्वार क्षेत्र हो स्टब्स्ट के स्टब्स्ट	
Arg426-Gly431B	(441)			क्षेत्रकार्याच्या ४ वर्गात्रकार वि	
Asn425-Lys432	(441)	्रद्राद्रद्रद्रद्रव्यक्ष्यात्वास	layer letter to the control of the	in inches de la	
Consensus	(441)	CCGCGGCGAGAT	CAAGAACTGCAGC	TTCAAGGTGACCACC	
		481		520	
Ile424-Ala433	(481)	יגבובן:יוונילפובישולינים	ोक्ताला है। इस्ताला के किस्ताला के किस	ilchrilge meise ser seit	
Trp427-Gly431				Hose fulcies contentions	
Gln422-Tyr435B		प्रदादरम् स्टब्स्ट्रास्ट्रास्ट्रास्ट्रास्ट्रा	initerial clear setting	संद्यानस्यत्यालनाकसम्बद्धाः	
Arg426-Gly431	(481)	nicionalitical contrata	sit-/cy:Andersite)talical	inchinistic clear week kind an	
Ile423-Met434	(481)	प्रदिक्तान्य द्वावी द्वारा ।	despectations for	Messivierie in inclusion	
Gln422-Tyr435	(481)	ACCOMPANIES CONTACTOR	क्षाप्त अस्तिक स्थापन । इ.स.च्या अस्तिक स्थापन	at supplied that the	
Arg426-Lys432	(481)	yelest freedning	sanding eligible tayou	Weberrele for ten Centre wit	
Arg426-Gly431B		decondesente	Market and expension and e	Meha:/@ic.en_@incirence?	
Asn425-Lys432	(481)			प्रकार रहेद्य १००० में क्षेत्र १००	
Consensus	(481)		AAGATGCAGAAGG	AGTACGCCCTGTTCT	
		521		5.60	
Ile424-Ala433	(521)	the thirt can be the factor	ske en action west a selection	or, respectively, or rejoint	
Trp427-Gly431	(521)	stoletate e telepare	inference of each profession	ભે ષ્ટ ાલ છ ાલમાં અન્ય વસ્તુ	
Gln422-Tyr435B	(521)	मुद्राममधीर गुन्तेश्वर्ता ह	म् स्वित्रमेद्रोद्देशक विभागतिक हैं।	经现代的 经现代的 医	
Arg426-Gly431	(521)	प्रकार एक विद्यान है।	हिल्लाम्बर्ग हे दिल्ला बाहरू	्रोत्रेक्ष्ण क्षेत्रका क्षेत्रको ज्याद	
Ile423-Met434		Activities deliver	itere verejernoù dieka	curried and principles	
Gln422-Tyr435	(521)	stabilities of electriciti	पुरुष्याचीत स्वाहरू पुरुष्य ।	新加州的特殊的	
Arg426-Lys432	(521)	dening a stable	देवां वा	计环心算形容别是现在	
Arg426-Gly431B	(521)	Herry Cardentel	प्रविद्यारातिक क्षेत्रप्रभिन्ता	NAME OF THE PROPERTY OF THE PROPERTY OF	
Asn425-Lys432	(521)	Activities adented	ificles nelesciolitics entit	characteristics of the confe	
Consensus	(521)		TGGTGCCCATCGA	CAACGACAACACCAG	
		561	·	600	
Ile424-Ala433				रास्त्रसायःभावसम्बद्धान्त	
Trp427-Gly431				detection is necessive	
Gln422-Tyr435B	(561)			Activities on the second of the	
Arg426-Gly431				non-left either eine eine eine E	
Ile423-Met434				ग्रह्मे स्टब्स्ट्रेस्ट्र्यूड स्टब्स्ट्रेस्ट्रिस्ट्रेस्ट्रिस्ट्रेस्ट्र	
Gln422-Tyr435		A TO STATE OF STATE OF	Report of Charles (Co.)	delicenter executive ente	
Arg426-Lys432				in the same elitities of the	
Arg426-Gly431B				र्याच्यासम्बद्धाः स्थाने स्थानिक वर्षात्	
Asn425-Lys432				defenderancial envirance	
Consensus	(561)		CAACTGCAACACC	AGCGTGATCACCCAG	
T1-404 T1 40-		601		640	
Ile424-Ala433				to:Ardmeethisedth	
Trp427-Gly431				द्वारम्बर्धद्वद्वारम् । द्वारम्बर्	
Gln422-Tyr435B				fediving elementaries editions in	
Arg426-Gly431	(601)	द्वेद्धाः स्वतं विवादाः । स्व	क्षा <u>त्कारदाक अपूर्</u> दक श्रम्	service distriction and the services of the se	
Ile423-Met434	(601)	ૡૡૼૡૡ૽ૡ૽ૢૺૡૡ૽ૼૢૼૡૢૡ૱ઌ૽૽૾ૡૢ	entransia in telephonesi	रेक्नेशनर १९४१ को प्यादश में बर्ग क्षेत्र र एक्षे <mark>त्र</mark>	

FIG. 4C

WO 00/39303		18 / 65	PCT/US99/31272
Gln422-Tyr435	(601)	वंद्यसम्बद्धियाः गृह्यसम्बद्धाः स्थाः व्यापद्धाः व्यापद्	esete(e)222te=te=te=t
Arg426-Lys432	(601)		
Arg426-Gly431B	(601)		
Asn425-Lys432	(601)		
Consensus		GCCTGCCCCAAGGTGAGCTTCGAGCCCAT	
	(002)	641	680
Ile424-Ala433	(641)	्रात्माद्यम्बद्धद्वद्वद्वद्वद्वद्वद्वद्वद्वद्वद्वद्वद्वद	
Trp427-Gly431	(641)		AVICE RULE TO DE LA
Gln422-Tyr435B		પ્રદાણના લાલા લાલા લાલા કરતા છે. તેને તાલા કરતા કરતા કરતા કરતા કરતા કરતા કરતા કરત	
Arg426-Gly431	(641)		
Ile423-Met434	(641)		
Gln422-Tyr435	(641)		
Arg426-Lys432	(641)		
Arg426-Gly431B	(641)		The state of the s
Asn425-Lys432	(641)		
Consensus	(641)	V	
consensus	(041)	681	720
Ile424-Ala433	16811	NATURAL PROPERTY OF STATE OF STREET	
Trp427-Gly431		a vier v Omiej v jelene velende en enselen	
Gln422-Tyr435B		entional physical redesice order	
Arg426-Gly431	(681)		
Ile423-Met434	(681)		
Gln422-Tyr435	(681)		
Arg426-Lys432	(681)		
Arg426-Gly431B	(681)		
Asn425-Lys432	(681)		
Consensus	(681)		
	•	721	760
Ile424-Ala433	(721)	TONE CONTRACTOR ACCOUNTS OF CONTRACTOR ACCOUNTS	cleyter terrent
Trp427-Gly431	(721)	Name had state and a control of the state of	લનાહલ્ય (લ) લાગુ
Gln422-Tyr435B	(721)	प्रकार-इद्राक्तिहर्मा स्थान स्थान स्थान स्थान स्थान	सक्ते र लक्ष र कार्य कार्य । इ.स.च्या १९८० विकास
Arg426-Gly431	(721)	यत्यसम्बद्धारम् इति । विद्याने । विद्याने । विद्याने ।	eerde eere
Ile423-Met434	(721)	येवर्गसम्बद्धम् स्ट्रोस्सम्बद्धारम् वस्यात् वस्यात् ।	
Gln422-Tyr435	(721)	ग्रेल्डिक्षेत्रस्तिरोज्ञाहरक्यानाः,दाप्रदेविताः एक्ष्यं सङ्गितिहाः	eteknytesinesidetekn
Arg426-Lys432	(721)	riciale is substancial consideration of the contemporary	
Arg426-Gly431B	(721)	्रहेल्लाहरू । अ प्रहा स्त्रहरूले । असिन्धार स्टब्स्ट्रेस	
Asn425-Lys432	(721)		
Consensus	(721)	ACCGTGCAGTGCACCCACGGCATCCGCCC	
71 404 71 477		761	800
Ile424-Ala433	(761)	चेर दर्भववाद्र तरुष्ट्रभम् स्थाप्तितास्य स्थापना स्थापन	
Trp427-Gly431	(761)		
Gln422-Tyr435B	(761)	elegization descriptions and inspection of the second description of t	
Arg426-Gly431	(761)	शतिकारेटः विशेषाति । इसकार स्ट्रिस्ट्रेस्ट्रेस्ट्रेस्ट्रेस्ट्रेस्ट्रेस्ट्रेस्ट्रेस्ट्रेस्ट्रेस्ट्रेस्ट्रेस्ट्र	
Ile423-Met434 Gln422-Tyr435	(761)	steroptoressensitietenstates meterostetens indensist	
Arg426-Lys432	(761) (761)	recommendations survivate and majorithese	
Arg426-Gly431B	(761)	क्षा अन्य विकास के प्रतिकार के अने का का किस्ता कर की अने का किस्ता कर की अने का किस्ता कर की अने का किस्ता कर	
Asn425-Lys432	(761)	જો <u>ાં પ્રાથમિક તેવામાં કોઇ તેવામાં પ્રાથમિક પ્રાથમિક</u>	
Consensus	(761)	CCCAGCTGCTGCTGAACGGCAGCCTGGCC	
Consensus	(101)	801	840
Ile424-Ala433	(801)	derminister (e. 16 e. 9.10.) To the Teach (e.)	The state of the s
Trp427-Gly431	(801)	der ayrı adayle ayları verili iye dayay	
Gln422-Tyr435B		covereded advantage and a	
Arg426-Gly431	(801)	चलप्रवासी संवर्षे प्रवर्षे प्रवास ्थलप्रकृतिक ।	
Ile423-Met434	(801)	deviatividelectideletionile;ic.://celetionile	
Gln422-Tyr435	(801)	वेद्यारकः भारतदास्त्रः । द्वाराम् । द्वाराम	
Arg426-Lys432	(801)	संस्कृतकारमण्डातकारकारकारमञ्जूष्टीस्थान	(Gerevitatice
· •		The same of the sa	

FIG. 4D

		19 / 03
Arg426-Gly431B	(801)	वर्गाचे संदर्भवद्यं वर्गन्ति । वर्ति । वर्गन्ति । वर्गन
Asn425-Lys432	(801)	वद्मारवः स्त्रज्ञल्यवः भवत् वर्षात् । भवत् । भवत् । भवत् । भवत् । भवत् । भवत् ।
Consensus	(801)	GGTGATCCGCAGCGAGACTTCACCGACAACGCCAAGACC
		841 880
Ile424-Ala433	(841)	ः प्रदेशकात्रकात्रकात्रकात्रकात्रकात्रकात्रकात्र
Trp427-Gly431	(841)	करले अधिक भवेते गये अधिक भिष्ठ प्राथित स्थापित स्थापित स्थापित स्थापित स्थापित स्थापित स्थापित स्थापित स्थापित
Gln422-Tyr435B	(841)	ॱग़ॺॣख़ऺॱढ़ढ़ख़॓क़ॿख़ॱॻॖख़॓ढ़फ़ॕॿ॓ऒख़ख़ऄॿऒख़ॱढ़ॿऒख़ॱढ़ढ़ऒॶॗ ॱ
Arg426-Gly431	(841)	ःत्रदृष्टेश्वरदृष्ट्यस्यक्षित्रस्यक्ष्मस्यक्षेत्रस्य । त्राचनस्य विकासम्बद्धाः । विकासम्बद्धाः । विकासम्बद्धाः
Ile423-Met434	(841)	भूरसम्पर्धाः प्रस्तिक स्वति । स
Gln422-Tyr435	(841)	. अर्थका भारत्ये भारत्ये भारत्ये अर्थका स्थापन । स्थापन स्थापन स्थापन स्थापन स्थापन स्थापन स्थापन स्थापन स्थापन
Arg426-Lys432	(841)	ામુદ્રકા, તુમ હલ્ફ કહા કરો મુક્ત કરો છે. તુમારા કરો
Arg426-Gly431B	(841)	
Asn425-Lys432	(841)	www.nkde.fidefile.fevil/legstel/levenkel/lefst/fest/fest/fest/fest/fest/fest/fest/
Consensus	(841)	ATCATCGTGCAGCTGAAGGAGGAGCGTGGAGATCAACTGCA 881 920
Ile424-Ala433	(881)	881 920 Send steer to the CANARA CONTROL AND
Trp427-Gly431	(881)	endelenantum in nerminen in
Gln422-Tyr435B	(881)	Sacidity of state the state of a state of size of size of size
Arg426-Gly431	(881)	sencing any and many and an analytic man and an and
Ile423-Met434	(881)	कें कर दिनिक हार कर कर के प्रतिक के कि कार कि
Gln422-Tyr435	(881)	સ્પાલદ (સંવહાર પ્રાંકારે પ્રહાર હોય સામાન કરે છે. જે છે. જે આ પ્રાંક કરે છે. જે આ પ્રાંક છે છે. જે માટે જે છે જે તાલદ (સંવહાર પ્રાંકારે પ્રહારે પ્રહાર હોય સામાન કરે છે. જે
Arg426-Lys432	(881)	estatetate il suprie convocanto, ne atere escontente productiva del tele
Arg426-Gly431B	(881)	jestogojojosavitos ir jesta jojos jestoja iz tra (ciut jih edija videk i isle
Asn425-Lys432	(881)	अवस्थितवर्षक्राम् द्राव्याम् स्थाने । त्रिकृतक्ष्याम् । त्रिकृतक्ष्याम् । त्रिकृतक्षयम् । व्यवस्थाने । वर्षकृत
Consensus	(881)	CCCGCCCCAACAACACCCCGCAAGAGCATCACCATCGG
		921 960
Ile424-Ala433	(921)	वकर्तिद्वादावद्वद्वद्वद्वद्वात्तरकाक्ष्यवद्वद्वतः । वद्वद्वद्वाद्वाद्वाद्वाद्वाद्वाद्वाद्वाद
Trp427-Gly431	(921)	જુદાલુકોલા તાલુકાલા કાલાકાર મારા છે. જે તાલુકાલા તાલુકા છે. તાલુકા છે. જે જે જે જે જે છે. જે જે જે જે જે જે જે જે જે જ
Gln422-Tyr435B	(921)	मान्य से सम्बद्धां त्रात्व व्यवस्थानम् कार्यः भ्रत्यात्र स्थानम् वर्षात्र स्थानम् अस्ति ।
Arg426-Gly431	(921)	श्रीतरात्रसम्बद्धाः स्थापना । स्थापना । स्थापना । स्थापना ।
Ile423-Met434	(921)	के खेलाड सम्बद्धित । यो सहित क्षेत्र का का स्वति सम्बद्धित । यो स्वति सम्बद्धित सम्बद्धित सम्बद्धित सम्बद्धित सम
Gln422-Tyr435	(921)	८०० चेल्यांनावन्त्रवाताः वर्षक्षम् (वेल्याम्) । जात्ये वर्षक्षम् । वर्षक्षम् । वर्षक्षम् ।
Arg426-Lys432	(921)	riferius teimeteminien in ministratizen error erenakteinistes et beväätetes e.
Arg426-Gly431B Asn425-Lys432	(921) (921)	ં વર્ષાના કર્યાલાલક કર્યા છે. કરાય છે. કરાય સામાના કરાય પ્રવાસના કરાય કરાય છે. સિંહાના વિવાસ કરિયાના કર્યા કર્યા છે. જે જો તેમ જો સામાના કર્યા કર્યા છે. જો જો સામાના કર્યા છે. જો જો જો જો જ
Consensus	(921)	CCCCGGCCGCCTTCTACGCCACCGGCGACATCATCGGC
Consensus	() 2 1 /	961 1000
Ile424-Ala433	(961)	avign socie stiller active reasons assent and designisting
Trp427-Gly431	(961)	क्षेत्रकात्रः अस्तिविक्षेत्रत्ताः विद्याराज्याविकार्त्रताः अस्ति स्वतिवित्तित्तरस्य
Gln422-Tyr435B	(961)	entrative electric entropic metric for a formal electric for the
Arg426-Gly431	(961)	केंग्रह अवतः हा इतिहास स्टाइंट क्षेत्र क्षेत्र का प्रदेश के तिस्त के स्टाइंट के क्षेत्र के क्षेत्र के क्षेत्र क
Ile423-Met434	(961)	anternation in Administration and the contraction and the contraction in
Gln422-Tyr435	(961)	रोशकार्यस्य संबंधितान् व्यवस्था । वेद अवस्थारं । अत्यासम्बद्धाः स्थान्यस्य
Arg426-Lys432	(961)	લ્લાના મામ્યાલી લાલ કાલના લાંભાર આવેલા છે. જે કાલના સાથે કરો કરો છે. તેને જે કાલના કાલના કરો છે. તેને જે કાલના
Arg426-Gly431B	(961)	PARTAIN SECRETARIOS RESISTINGES DAS RECURSORIOS CICIONES AAS DE
Asn425-Lys432	(961)	arter the contraction one garden activities and activities activities and activities and activities activities activities and activities acti
Consensus	(961)	GACATCCGCCAGGCCCACTGCAACATCAGCGGCGAGAAGT
		1001 1040
Ile424-Ala433	(1001)	elentarinitation in interpretation and electric contraction of the con
Trp427-Gly431	(1001)	eleftat/cita.ccst/ccelenteatat/ele/teathacheles/ele/ascitation-taticion
Gln422-Tyr435B	(1001)	वंक्रकोत्राकोत्रात्तात्त्वके क्ष्मकार्यस्य । त्याच्यात्यात्यात्यात्यात्यात्यात्यात्यात्यात
Arg426-Gly431	(1001)	elegyatetteinetatetetatateten lettpastejantatroottattelen netejatete
Ile423-Met434	(1001)	वस्यक्षात्रम् । त्यान्यक्षात्रम् । त्यान्यक्षात्रम् । त्यान्यक्षात्रम् । त्यान्यक्षात्रम् । त्यान्यक्षात्रम् ।
Gln422-Tyr435	(1001)	लेक्षोक्रमक्षेत्रम् स्वतः क्षेत्रक्षेत्रम् । इत्योक्षेत्रम् । इत्योक्षेत्रम् । इत्योक्षेत्रम् । इत्योक्षेत्रम्
Arg426-Lys432	(1001)	द्धां प्रशेषां भगवतं संस्थात् वर्षे प्रवेषाः भगवतः प्रशेषात् । भगवतः प्रशेषात् । भगवतः भगवतः ।
Arg426-Gly431B Asn425-Lys432	(1001) (1001)	etrivicavimi er relavido de legente de la legente de l
valla52_DA8425	(1001)	व्यवस्थितः । । । । । । । । । । । । । । । । । । ।

FIG. 4E

Consensus (1001) GGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGC 1041 1080 इतिमित्रमान्यविद्यम् । लाक्षेत्रमान्यविद्यम् । विद्यान्यविद्यम् Ile424-Ala433 (1041) (1041) व्हार्यक्ष्मिक्ष्मिक्ष्मित्र्वावक्ष्मिक्ष्मिक्ष्मिक्ष्मिक्ष्मिक्ष्मिक्ष्मिक्ष्मिक्ष Trp427-Gly431 Gln422-Tyr435B (1041) Actaenhamicielehandenhamikenantenhamikenandenhamicielehan (1041) इंदेम्ब्यूम्प्रियंस्याद्वास्त्रक्रम् व्याप्तिक स्ट्रांस्य स्ट्रांस्य स्ट्रांस्य स्ट्रांस्य स्ट्रांस्य Arq426-Gly431 Ile423-Met434 (1041) विकास सामित्र विकास स्थानिक स्थ Gln422-Tyr435 (1041) व्यवस्थानिक विकास कार्या होता विकास के अन्तर कार्या Arg426-Lys432 (1041) menterbitomum patomater metalicie del freche automater acte Arg426-Gly431B (1041) चेकालामा संस्थान क्षेत्रका स्थापन स्यापन स्थापन स्यापन स्थापन स्य (1041) FOR CONTROL OF THE ANGEL OF THE PROPERTY AND PROPE Asn425-Lys432 Consensus 1081 Ile424-Ala433 (1081) लेसन्यल्डाकरानसञ्जलसरोक्तराहरू स्वरूपसङ्ख्याकरासील नार्व Trp427-Gly431 Gln422-Tyr435B Arg426-Gly431 चेद्रांद्रचित्रदर्भात्रद्रचेत्रेत्रम्भात्रद्रविद्रम्भावन्त्रम्भात्रद्रभावन्त्रम्भात्। Ile423-Met434 (1081) Gln422-Tyr435 (1081) बद्धार्यकारकार्यक्षित्रविद्धारम्भद्दारकार्याद्वार्थारकार्याद्वारकार्याद्वारकार्याद्वारकार्याद्वारकार्याद्वारक Arg426-Lys432 (1081) gdaddrawdelederid xederid weder one en richt gebede (1081) geoletestatuleleidettestatelesteinamentettestatuleide Arg426-Gly431B Asn425-Lys432 (1081) इंदर विद्यापक व्यवस्थित व्यवस्थित विद्यापक व्यवस्था व्यवस्थित विद्यापक विद्या Consensus (1081) GGCGGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCG 1121 enterested templeten entitlen fellen fellen han beste entereste en enter fellen en en Ile424-Ala433 (1121)Trp427-Gly431 (1121)द्विद्वात्त्वभूष्यापुरुषा श्रेष्ट्राण्ये स्थाप्तात् । स्थाप्तात् भूति स्थाप्ता । Gln422-Tyr435B (1121)विवादीवाहरू प्रकाशिक विवेद कर १८०१ विकास प्रकाशिक प्रकाशिक विवाद है। Arg426-Glv431 (1121) मृत्यं कर्ता । विकास क्षां क्षा Ile423-Met434 Gln422-Tyr435 (1121)દાં સાંદર્ભાદા ભાગમાં માર્ગા છે. જે કાં તાલા મુખ્ય અને કાં મારા છે. આ પ્રાથમિક સ્ટેપ્ Arg426-Lys432 (1121)Arg426-Glv431B (1121) an activity feather and chills in each contact the activities will be a contact the activity of Asn425-Lys432 desirence yeneral annitante of the second se (1121)Consensus (1121) GCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAA Ile424-Ala433 (1161) Resemble to the state of Trp427-Gly431 (1161) AND CONTRACTOR OF THE C Gln422-Tyr435B (1161) Arg426-Gly431 (1161) Process a consideration of the concentration of the constant of the con Ile423-Met434 (1161) where the construction of the construction of the construction of Gln422-Tyr435 (1161) where the experience of the property of the second Arg426-Lys432 (1161)challes also matelesias minales y chalancieles a constitue de la citation de la citation de la company de la compa Arg426-Gly431B attentional connection of the language and a contribution of the section of the s (1161)Asn425-Lys432 (1161) Masservers masservers was experienced as a confidence of the confidence of th Consensus CAGCACCTGGAACAACACCATCGGCCCCAACAACACCAAC (1161)1201 Ile424-Ala433 (1201)विद्धाप्रदासम्पूर्णकार्थाः क्षार्वकारविद्यात्रेष्ठेत्रकारम् । स्थानिकार्यकारम् । Trp427-Gly431 (1201)decention to second acondition to the contraction of the contraction o (1201) SECONO MENGRACION CON CONTROL MANAGERO CONTROL MAN Gln422-Tyr435B Arg426-Gly431 (1201) व्यवस्थानाम् अस्य स्थानिक विकास स्थानिक विकास स्थानिक स Ile423-Met434 (1201) Selectic coveres consists of the Control of Gln422-Tyr435 (1201) deistre in the state of Arg426-Lys432 (1201) deletative deletative deletative del constructive Arg426-Gly431B (1201) विकासिक प्रकारिक स्थानिक स्थानि Asn425-Lys432 (1201) delotación del especial del control de la control d Consensus (1201) GGCACCATCACCCTGCCCTGCCGCATCAAGCAGATCATCA 1241 1280

FIG. 4F

PCT/US99/31272

FIG. 4G

Gln422-Tyr435B	(1417)	ब्रह्मेयर तहेया वर्षे वर्षे हें के द्वारा वर्षे
Arg426-Gly431	(1441)	ૡૺૡ૽ૡ૽૱ૹૡૡૡ૽ૡૡ૽ૡૢૡ૽ૡ૽ૡ૽૽ઌ૽૽ૡ૱ૡ૽ઌ૾૱ૹૡૡૡઌૹૹૡ૽ૡૡૡૡૡૡૡૡ
Ile423-Met434	(1423)	न्तिक न्यंब्राचा स्वतं व्यवस्थित । स्वतं व्यवस्थित स्वतं व्यवस्थ
Gln422-Tyr435	(1417)	सर्वहाः महास्वयहार्द्धस्यविद्यातिकृतिहार् अवद्यविद्यान्त्रविद्यान्त्रविद्यान्त्रविद्
Arg426-Lys432	(1441)	च्य (या भ्रेस्ट्रोलेटा तम्ब्रेसाच्याचा वाचा वाचा प्रकार का भ्याचा वाचा हो भूते वाचा वाचा वाचा वाचा वाचा वाचा व
Arg426-Gly431B	(1441)	न्तर्भारः वृद्दिद्वाः होकान्द्रद्वित्राद्वित्राद्वाराम् । वृद्धान्त्राद्वाराम् । वृद्धान्त्राद्वार्थे ।
Asn425-Lys432	(1435)	द्वयं व्याप्त विद्या के साम के त्राप्त के त
Consensus	(1441)	CCCCTGGGCGTGGCCCCCACCAAGGCCAAGCGCCGCGTGG
		1481 1520
Ile424-Ala433	(1469)	Redainstance to the city of cate of the standard confirmation
Trp427-Gly431	(1481)	अवकारदेखंबदेविपुरुक्तां खालबाबुद्धां सारकारदाव्यां प्रवादां द्विता हुए। संवकारदेखंबदेविपुरुक्तां खालबाबुद्धां सारकारदाव्यां प्रवादां द्विता हुए।
Gln422-Tyr435B	(1457)	भवक्तावल्यां क्षितं क्षारे मा कार्य कार्य कार्य क्षेत्र स्वतं विकास क्षेत्र स्व
Arg426-Gly431	(1481)	त्रहात्रमद्वाद्वाद्वाम् मृत्युव्याद्वाद्वाद्वाद्वाद्वात्वात्वात्वात्वात्वात्वात्वात्वात्वात
Ile423-Met434	(1463)	ाः होद्रम्भाने लेखाः सम्प्राप्ताके ने हासम्बद्धाः स्थापः सम्प्राप्ताः । । । । । । । । । । । । । । । । । । ।
Gln422-Tyr435	(1457)	તારીએમાં લાકાલ માલોક મદાઉદાદી લાકા જેક કરાય લાકાલા લાકાલા વાર્ય કરો છે. ઉદ્યાનો
Arg426-Lys432	(1481)	ाम्यांकोत्राच्यांच्यांकांकांकांकांकांकांकांकांक्यांच्यांकांकांकांक्यांकांक्यांच्यांच्यांच्यांच्यांच्यां
Arg426-Gly431B	(1481)	१ ८९ ५१८७६५४५५४५५५५५५५५५५५५५५५५५५५५५५५५५५५
Asn425-Lys432	(1475)	एसको स्टब्स्ट्रास्ट्रास्ट्रास्ट्रास्ट्रास्ट्रास्ट्रास्ट्रास्ट्रास्ट्रास्ट्रास्ट्रास्ट्रास्ट्रास्ट्रास्ट्रास्ट्
Consensus	(1481)	TGCAGCGCGAGAAGCGCGCCGTGACCCTGGGCGCCATGTT
•	\	1521 1560
Ile424-Ala433	(1509)	स्वत्रमदासम्बद्धाः महत्त्वम् । संबद्धाः स्वतं ।
Trp427-Gly431	(1521)	भेटोमद्दित्व वाद्याचेत्रावृत्ति हेटात्राचेत्रात्त्रात्त्रात्त्रात्त्रात्त्रात्त्रात्त्रात्त्रात्त्रात्त्र
Gln422-Tyr435B	(1497)	ૹૹૡ૽૽ૡ૽ૡૺ૱૽ઌૺૹૡૡ૽૽ઌૡૺૺઌૺૡૡૺૡઌ૽ૡઌૡૹૡૹૡૹૡૢૡૡૡૡૡૡ <mark>ૺ</mark>
Arg426-Gly431	(1521)	ाकार का व्यवस्था में का
Ile423-Met434	(1503)	मुक्तिदादावर्गमध्यस्य स्वातंत्रवस्य स्वतंत्रस्य स्वतंत्रस्य स्वतंत्रस्य स्वतंत्रस्य स्वतंत्रस्य स्वतंत्रस्य स
Gln422-Tyr435	(1497)	ॳॖक़ग़ढ़ऻढ़ॏढ़ॱॿॱॿढ़ज़ढ़ग़ ਜ਼ਜ਼ਫ਼ੑਜ਼ ਲ਼ਫ਼ਜ਼ਫ਼ੑਖ਼ਖ਼ੑਖ਼ਖ਼ੑਖ਼ੑਜ਼ੑਖ਼ੑਜ਼ਲ਼ਖ਼ਫ਼ਲ਼ਖ਼ਜ਼ਫ਼ਜ਼ਜ਼ਜ਼ਜ਼
Arg426-Lys432	(1521)	ीक १९ दर्शन १५ १५ १५ विद्यालयम् । १५ १५ १५ १५ १५ १५ १५ १५ १५ १५ १५ १५
Arg426-Gly431B	(1521)	क्रमान्द्रद्वात्त्रद्वात्त्रद्वात्त्रं क्षाचनन्त्रं क्षाच्यात्त्रं स्टब्स्ट्रेट स्टब्स्ट्रेस्ट्रेस्ट्रेस्ट्रेस्
Asn425-Lys432	(1515)	प्रदेश, वार्षाच्या क्षेत्र क्षेत्र क्षेत्र वार्षाच्या क्षेत्र क्षेत्र क्षेत्र क्षेत्र क्षेत्र क्षेत्र क्षेत्र क
Consensus	(1521)	CCTGGGCTTCCTGGGCGCCGCCGGCAGCACCATGGGCGCC
	(1011)	1561 1600
Ile424-Ala433	(1549)	्रवातामान्यः । क्षेत्रवाद्द्राक्षः चत्रदेशकान्यः । कृत्यः चत्रकान्यः । वर्षः । वर्षः । वर्षः । वर्षः । वर्षः ।
Trp427-Gly431	(1561)	પુંચલમાં કુલ કુલ જાણ જાણ કુલ જાણ છે. જો છે
Gln422-Tyr435B	(1537)	strander som mentionades accordenates and
Arg426-Gly431	(1561)	elassinista katis elimitriti past jastilaala kijakinist kiiaentsti
Ile423-Met434	(1543)	Surviva with the Act of early the survival detail of the Colored Section 19
Gln422-Tyr435	(1537)	लंबांबर्गचरारः । असः स्वरं द्रवस्थात्वर्षे कार्यात्रेवस्य व्यक्तिस्य वर्षे
Arg426-Lys432	(1561)	aldered desired enterprise and a semile of the enterprise and
Arg426-Gly431B	(1561)	etelesteleleneatrioleneatrolenelenelelelelelelenelenelenelenelenel
Asn425-Lys432	(1555)	्रातास्त्रात्वा अवस्थितस्य स्थानस्य स्थानस्य स्थानस्य स्थानस्य स्थानस्य स्थानस्य स्थानस्य स्थानस्य स्थानस्य स्
Consensus	(1561)	CGCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGA
Consensus	(1301)	1601 1640
Ile424-Ala433	(1589)	दावादावाकार व्यवस्था हो स्टब्स्स स्थापना स्थापना स्थापना स्थापना स्थापना स्थापना स्थापना स्थापना स्थापना स्थापन
Trp427-Gly431	(1601)	
Gln422-Tyr435B	(1577)	za zada nanjete na 20 frejetarete prozaka oporten estakoj indeze tre ost Egiz ek 1999 a prozak testota poprava obska objeti i stakoje tipete i ilij
Arg426-Gly431	(1601)	
Ile423-Met434	(1583)	क्षण होता को स्पर्धित क्षित्र कि क्षण के हिंद कर है। इस क्षण होते हैं कि क्षण होता है। इस होता है। विकास के स्पर्धित क्षण के कि
Gln422-Tyr435	(1577)	gloesta madeente enterratura abancan en entere da m
Arg426-Lys432	(1601)	કુંકુ(લાનાના કુકુલકુમલા અનુક દુકુલાનું કુકુલા કુકુલાનું કુકુલાનું કુકુલાનું કુકુલાનું કુકુલાનું કુકુલાનું કુકુ ત્રામ
Arg426-Gly431B		દાં મુન્યમાં ભારતે કહોતા-જોલો માતાનો દામની કર્માં ભારત મહોલોલો છેલા છે. જો
Asn425-Lys432	(1601) (1595)	चन्त्रकाराज्यकारमञ्जूष्यकारम् । जन्मकाराज्यकारम् । जन्मकाराज्यकारम् । जन्मकाराज्यकारम् । जन्मकाराज्यकारम् । जन
Consensus		CCCCCATCCTCCACCACCACCACCACCACCACCACCACCA
Consensus	(1601)	GCGGCATCGTGCAGCAGCAGAACAACCTGCTGCGCGCCAT
T16424-11-422	(1620)	1641 1680
Ile424-Ala433 Trp427-Gly431	(1629)	स्याग्न्यपुर्वेष्टासम्बद्धान्यम् स्वत्यस्यकार्यकार्यम् । स्वाग्न्यपुर्वेष्टासम्बद्धान्यम् ।
Gln422-Tyr435B		enentrales serviciente contralenta del necessa de presidente se de contrales de la contrale de la contrale de contrales de la
		ंतिराहोत्तरः अरोत्तवरात्। जगराः रत्निमात्तर्भात्त्वः अति श्लितिमानि अरिति ।
Arg426-Gly431	(1641)	ल्लारवित्रविद्धारम्बाकारव व्यक्तः । विवाद अवाव स्थान व्यक्तरकः व्यक्तरकः व्यक्तराम्

FIG. 4H

Ile423-Met434	(1623)	तंदान्यस्तराज्ञस्यस्त्रेत्रस्याराज्ञस्यम्
Gln422-Tyr435	(1617)	
Arg426-Lys432	(1641)	eretatelerereterateleriaterateraterateraterateratatatatatatatat
Arg426-Gly431B	(1641)	संदेशवासामानामानामानामानामानामानामानामानामानाम
Asn425-Lys432	(1635)	नवार्यात्रात्रात्रात्रात्रात्रात्रात्रात्रात्र
Consensus	(1641)	CGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGGC
	(2012)	1681 1720
Ile424-Ala433	(1669)	ingerviscillelendervelletigesjendomaldesporatifieserif
Trp427-Gly431	(1681)	भूदर्गभूमा । ज्याने प्रकार प्रवास का स्थापन । ज्याने का स्थापन । ज्याने । ज्याने । ज्याने । ज्याने । ज्याने ।
Gln422-Tyr435B	(1657)	भारक्षात्रभेत् कामाराज्येद्वेत्यक्षत्त्रभ्यक्षत्त्रभ्यक्षत्त्रभयद्वेत्रभयद्वेत
Arg426-Gly431	(1681)	/www.vec.steams.come.come.come.come.come.com
Ile423-Met434	(1663)	
Gln422-Tyr435	(1657)	ૹૢઌ૽ૡઌૡઌૡઌૡઌ૽ૡૢઌ૽૱ૡૡ૽૽ઌૡૹૢઌઌૡૢઌ૽ૺ૾૽૱ૹઌૡૡ ૡ ઌૡઌૡ
Arg426-Lys432	(1681)	का विश्व राजनायकः जनसम्बद्धारम् विद्वारम् म्हणसम्बद्धारम् स्तर्भन्ति स्तर्
		क्षित्रकार्वाद्वाकार्वाद्वाक्षयकार्वाद्याच्याक्षयक्ष्यच्याच्याच्याक्षयकार्वाद्याच्या
Arg426-Gly431B	(1681)	<i>ૢૡ</i> ૽૱ઌૡઌ૽ઌૢૡૢૡૡૡૡઌ૽ૡઌ૽ઌૢૡૢૡઌૢઌઌઙઌ૽૽૽ૢૡૡૡઌૡઌૢૡઌૢઌૢઌૢઌ૱૱
Asn425-Lys432	(1675)	कान्त्रकार विकास देव स्थापित हो हो हो है । इस का अधीर की हो से स्थाप हो हो हो ।
Consensus	(1681)	ATCAAGCAGCTGCAGGCCCGCGTGCTGGAGCGCT
73 404 33 433		1721 1760
Ile424-Ala433	(1709)	रात्वारम्भवारम् रवादाः मृत्यार्थात् । रद्यार्थः रद्याः सम्बद्धाः स्वतावत् विद्यान् । सम्बद्धाः स्व
Trp427-Gly431	(1721)	प्रदेश महाभागवद्यम् । अप्रदेशमञ्जानम् । वद्यानाम् । वद्यानाम् । वद्यानाम् । वद्यानाम् ।
Gln422-Tyr435B	(1697)	प्रश्रदेश वाप्त्रमा विभागति पृष्टीकामा आक्रमेश वास्त्रिका विश्वप्रदेश विभाग स्थापन स्थ
Arg426-Gly431	(1721)	yele, new major established major medanjing meredere ner Ng
Ile423-Met434	(1703)	पूर्वकामधारम्भित्रेर्वात्रारम्भाग्यामभ्यतिकृतिसम्बद्धारम्भाग्यतिस्तित्वित्वित्वित्
Gln422-Tyr435	(1697)	र्गत्त्रशास्त्रभावत्त्रशास्त्रवाद्वत्त्रभावत्त्रभावत्त्रभावत्त्रभावत्त्रभावत्त्रभावत्त्रभावत्त्रभ
Arg426-Lys432	(1721)	Fig. no. Avelogy with classic degrees and contraction of the
Arg426-Gly431B	(1721)	યું કુલ્યું કરા કર્યો છે. તે કોલ્યું કોલ્યું કે કે કે કે કુલ્યું કરે કે કુલ્યું કે કે કુલ્યું કે કે કુલ્યું કે
Asn425-Lys432	(1715)	Apriliabilitical swell do land described in sold elements.
Consensus	(1721)	ACCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCAG
T1-424 N1-422	/17/01	1761 1800
Ile424-Ala433	(1749)	relativization sanda venivarial delateración estadivida e
Trp427-Gly431	(1761)	eliteres y a comercial contratat de constante es a describir con sector de la contrata del contrata del contrata de la contrata del la contrata de la contrata del contrata del la
Gln422-Tyr435B	(1737)	elefologiation of the spiritual state of the spiritual section of the s
Arg426-Gly431	(1761)	द्रोदांकर करान्त्र । क्यांकांकां का व्यवस्थित दिविभेदादी हार प्रभावने क्षेत्रक विकास भी स्थापन भी स्थापन
Ile423-Met434	(1743)	Reportation (16)) - Series of Autorition (16) decided a glassical and the
Gln422-Tyr435	(1737)	ા તુલાનું કહેર કર્યું . પુંચલ પાનું અને ભાગમાં ભાગમાં તુલાનું પ્રત્યાન છે છે. જે
Arg426-Lys432	(1761)	Delegativiter til krives og fill til elektrike elektriket elektrik
Arg426-Gly431B	(1761)	sidele paren premium process of folelogo de diseputado a tradicio.
Asn425-Lys432	(1755)	हे होहों । अर्थमा होते हेर्स हेर्स हरा कार्य को स्टोल्स हो होते हैं से क्षा होता है
Consensus	(1761)	CGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCC
Ile424-Ala433	(1700)	1801 1840
Trp427-Gly431	(1789)	રોલ : માં, લોલો રોલોલો કરે કરો કેટ કરો છે. કરો લોલો લો કરો કરો કરો કરો કરો છે. કરો કરો કરો કરો કરો કરો કરો કરો
Gln422-Tyr435B	(1801)	Helesseles in Assumantantanean and the second and the second
Arg426-Gly431	(1777)	Artigate white the contraction of the contraction o
Ile423-Met434	(1801)	ticks has strategic to propositionaries reactive to the classic sistems.
Gln422-Tyr435	(1783)	their terminated three religions are the principle strains in the
Arg426-Lys432	(1777)	Nate warmen mit manifyan fantium e in elektric om tenne sigen elektrick bydest.
Arg426-Gly431B	(1801)	neles sejeintelestratetinferneren sejeintelenterannen seleintelestraties
Asn425-Lys432	(1801)	ne emparatelante nationi de materiales recentorismente partenate de materiales de mate
Consensus	(1795) (1801)	Actual value
Consensus	(1001)	AGCTGGAGCAACAAGAGCCTGGACCAGATCTGGAACAACA 1841 1880
Ile424-Ala433	(1829)	hairogreathradamadamannathradadamana
Trp427-Gly431	(1841)	#CANCERSON ESTACTED COASCISSON CARGOAS ACCUMENCE
Gln422-Tyr435B	(1817)	delygen descaded teadels adeletated participation and the production of the contract of the co
Arg426-Gly431		त्रदार्श्वरात्रसंद्रकात्राहरूकात्रसंद्राहरूतात्वराद्रीहात्। स्वतंत्राहरूतात्वरात्राहरूतावर्गात्वरात्
Ile423-Met434		ignizationalmineum sepreteum puramente i septimani pariteitiste te
Gln422-Tyr435	(1817)	neinterspheisterner are seiten promit in der eine in halte helde eine
•		FIC AT

FIG. 4I

Arg426-Lys432	(1841)	ः स्वतः (१९७४:१९) वर्षः स्वतः ।
Arg426-Gly431B	(1841)	ः। वर्षितव्याम् स्वतं मेमस्य वर्षिते सद्देवत् वर्षित् वर्षित् । वर्षित् वर्षित् । वर्षित् । वर्षित् । वर्षित् ।
Asn425-Lys432	(1835)	इत्याद्वाक्त्रवाद्वीकृत्वकृतिकात्रवाद्वात्रीत्वाद्वात्रकृतकात्वाद्वात्रकृतिकात्
Consensus	(1841)	TGACCTGGATGGAGTGGGAGCGCGAGATCGACAACTACAC
	••	1881 1920
Ile424-Ala433	(1869)	योग्नीकोद्वाना,द्वाताखाः(श्राद्यक्षयानेत्र्यखांच्यायानाच्यकान्याकाराहाः।
Trp427-Gly431	(1881)	ระหวัดเล่นสมรัชย์เร็บสำคัญสมรัสเดิมเล่น เล่น ของเล่น เล่น เล่น เล่น เล่น เล่น เล่น เล่น
Gln422-Tyr435B	(1857)	लेक्प्रकार, द्वाकार एकं एक प्रदेश होता के काल गिरिक्त ग्रांचित प्रकार होते हैं। विकास का प्रदेश का का का प्रकार
Arg426-Gly431	(1881)	estates contain to applicated by a graph of estates the free training the safe
Ile423-Met434	(1863)	erantelengengengennappteletengenmentaliselendese benyere en togskiften de
Gln422-Tyr435	(1857)	ajtatyche notyknejatych folganiste i safatytejetatytejetaj se svijenie ve
Arg426-Lys432	(1881)	efections and the engine of the months of content of the introduction of the content of the cont
Arg426-Gly431B	(1881)	चारत्रेत्रे इत्रेष्ट्रे स्थापन कर्षेत्र के त्रेष्ट्र का स्थापन स्थापन स्थापन स्थापन स्थापन स्थापन स्थापन स्थापन
Asn425-Lys432	(1875)	ब्राक्षांद्रातालुक्तां वर्षा क्षां क्षां क्षां व्यवद्वाता क्षां क्षां वर्षा क्षां वर्षा क्षां वर्षा क्षां द्वा
Consensus	(1881)	CAACCTGATCTACACCCTGATCGAGGAGGCCAGAACCAG
		1921 1960
Ile424-Ala433	(1909)	का व्यक्त महामा (विकास महामा का अपने का स्वर महिल्ला है। असे का स्वर महिल्ला का स्वर महिल्ला का स्वर महिल्ला क
Trp427-Gly431	(1921)	એ પ્રાંક મુંદ્રો કે માનું કે મુક્તિ પાસ્તુ છે. પ્રાંક મુક્તિ માનું મુક્તિ માનું મુક્તિ માનું મુક્તિ માનું મુક્ માનું મુક્તિ માનું
Gln422-Tyr435B	(1897)	entagnonne angandoriaevalence addition comitain val
Arg426-Gly431	(1921)	citienadesatumastajnistojestajaintojnimosneteides mideidetajdies
Ile423-Met434	(1903)	signifestionistytestyteistyteistyteistys super meteryclomateristyteis
Gln422-Tyr435	(1897)	while delications are estimated and a federic accountance
Arg426-Lys432	(1921)	કાર્યાલમાં માત્ર કર્યા હતા. કર્યા હતા. કર્યા હતા. કર્યા હતા. કર્યા હતા. કર્યા હતા. કર્યા છે કર્યા છે છે છે છે.
Arg426-Gly431B	(1921)	અનુતાનુ પ્રાથમ કર્યા હતું હતું છે. અમાનું લેખે લેખ મુંદદ હતા છે. જે
Asn425-Lys432	(1915)	ामान्यात्राक्षेत्रं स्थाप्यात्रात्राच्यात्रात्रात्रात्रात्रात्रात्रात्रात्रात्र
Consensus	(1921)	CAGGAGAAGAACGAGCAGGAGCTGCTGGACCAAGT
	,,	1961 2000
Ile424-Ala433	(1949)	digaster/statestentatestricasatentisates estructutatestricastesting
Trp427-Gly431	(1961)	लेक्स्य अस्तर्भक्षा स्थापने व्यवस्था । अस्तर्भक्षा अस्ति । अस्
Gln422-Tyr435B	(1937)	नंदर्भः अवन् वर्षान्तं । वर्षान्तं । वर्षान्तं ।
Arg426-Gly431	(1961)	सङ्ग्रहारम् । तत्त्रकारमञ्जूषाम् । । । । । । । । । । । । । । । । । । ।
Ile423-Met434	(1943)	स्ट्रिस्टार्गात्र्वर्गात्राच्या स्ट्राक्षेत्रम् । सन्दर्भात्रः सम्बद्धाः वर्गत्रः वर्गत्रस्थात्रम् स्ट्राप्तस्
Gln422-Tyr435	(1937)	राम्यावार । व्यवस्था । वाद्या सारा व्यवस्था । सामान्या । सामान्या । सामान्या । सामान्या । सामान्या । सामान्या
Arg426-Lys432	(1961)	वेदी देदीया. राजेंद्री ते पूर्वमा वाद्या १६० स्वयंत्र १६० १६० १६० १६० १६० १६० १६० १६० १६० १६०
Arg426-Gly431B	(1961)	didentify de mentier au de mande de man
Asn425-Lys432	(1955)	Bottlerande belongrapherena etgagian hacabate kelengi
Consensus	(1961)	GGGCCAGCCTGTGGAACTGGTTCGACATCAGCAAGTGGCT
	(2002)	2001 2040
Ile424-Ala433	(1989)	त्वत्रात्वस्य विशेषात्रा प्रतिकृतिक स्वतं विशेषात्र विशेषात्र स्वतं त्या स्वतं विशेषात्र स्वतं विशेषात्र स्वतं
Trp427-Gly431	(2001)	के तेतृह भग्ना का
Gln422-Tyr435B	(1977)	chicles are a strategic and through a to so and in a characteristic se
Arg426-Gly431	(2001)	क्षांबरद्रमानसम्बद्धाः प्रभावन्त्रस्य स्थानि ।
Ile423-Met434	(1983)	the separate residence of selections and the second control of the
Gln422-Tyr435	(1977)	aparatoj, nivejas a prisado izmirajnih alistija iz izmiras na maju prisade iz izlete iz iniĝ
Arg426-Lys432	(2001)	sandon program director in the entering of the control of the city
Arg426-Gly431B	(2001)	areachaneaneasadenearpaneareachaneareachan cald clacke ee
Asn425-Lys432	(1995)	chnoleinhiteranathiteinhine ann cinn chaidh an chnollochtich a'c
Consensus	(2001)	GTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTG
	(2001)	2041 2080
Ile424-Ala433	(2029)	antandemental and principal and the state of
Trp427-Gly431	(2041)	chileterne recietablice soldin exterein rementere encience.
Gln422-Tyr435B	(2017)	ಪ್ರಕರ್ಷಕರು ರಾಕ್ಷರ ಹಲದು ಕುರೇವಕ್ಕಿಯಾಗಿಕೆಯಲ್ಲಿ ಅಗೆ ಕಾರಿಗೆ ಗಾಗಿ ಬಿಡುಗಳು ಅಗಳ ಗರುತ್ತಿ
Arg426-Gly431	(2041)	कारवादाः कारवादाः भागवादावाद्याः विश्ववादान्ति । विद्यानाः विद्यानाः
Ile423-Met434	(2023)	examencial original adaption of the control of the control or of the control of t
Gln422-Tyr435	(2017)	enalitient antiter august neis greicht fer Vertrember auf der Vertrember der Vertrempfer der
Arg426-Lys432	(2041)	chacle of the celestric confidence and confidence to the confidence of
Arg426-Gly431B	(2041)	eksteleleitigats, ein aastetamen tin etenheiten ei trocktoneiten (en taket eine en trocktoneiten (en taket eine En alleten ein ein ein er viert betrauten en en er en alleten (en taket eines eines en de staket en trocktonei
J	,/	EIC A I

FIG. 4J

	10005	N == 435 1 430	
	(2035)	Asn425-Lys432	
	(2041)	Consensus	
	(2069)	Ile424-Ala433	
	(2081)	Trp427-Gly431	
) प्रवस्तिकारवेद्यस् प्रवस्तिकारकारः स्टारविकारम् विकासिकार्	(2057)	Gln422-Tyr435B	
	(2081)	Arg426-Gly431	
	(2063)	Ile423-Met434	
) antidated gateletratele artisticate etalia per per primitario etalia.	(2057)	Gln422-Tyr435	
	(2081)	Arg426-Lys432	
	(2081)	Arg426-Gly431B	
	(2075)	Asn425-Lys432	
	(2081)	Consensus	
2121 2160		•	
) व्यवस्थाः इतिहास्यक्षेत्रः स्वर्धस्य स्वतः स्वतः स्वतः स्वतः स्वयं स्वयं स्वयं स्वयं स्वयं स्वयं स्वयं स्वयं	(2109)	Ile424-Ala433	
्रवित्रद्वाः सर्भवत्रवेद्वेत्रव्यवस्त्रव्यव्यवस्त्रव्यव्यवस्त्रव्यवस्त्रव्यवस्त्रव्यवस्त्रव्यवस्त्रव्यवस्त्रव्य भाषाम्	(2121)	Trp427-Gly431	
	(2097)	Gln422-Tyr435B	
	(2121)	Arg426-Gly431	
	(2103)	Ile423-Met434	
	(2097)	Gln422-Tyr435	
	(2121)	Arg426-Lys432	
	(2121)	Arg426-Gly431B	
	(2115)	Asn425-Lys432	
0.64	(2121)	Consensus	
2161 2200	(01.40)	Ile424-Ala433	
	(2149)		
	(2161)	Trp427-Gly431 Gln422-Tyr435B	
The state of the s	(2137) (2161)	Arg426-Gly431	
	(2143)	Ile423-Met434	
special and additional coloridates of the section o	(2137)	Gln422-Tyr435	
	(2161)	Arg426-Lys432	
જ્યાં વ્યાપાલ વાલું મેં કહેલું એ લેક કરો છે. સંસ્થિત સામાન	(2161)	Arg426-Gly431B	
	(2155)	Asn425-Lys432	
	(2161)	Consensus	
2201 2240	(2202)		
	(2189)	Ile424-Ala433	
	(2201)	Trp427-Gly431	
नेहर्मित्रात्वः वर्षे दृष्ट्यायान्यः । अवश्यात्वात्वात्यः । अस्य ।	(2177)	Gln422-Tyr435B	
che fotore e and consolerante for our party areas are a community content of the	(2201)	Arg426-Gly431	
eleteletetakki etkeletite kieloje netrinalatiin ekipaliti quiseteletit kiisi	(2183)	Ile423-Met434	
etero estrumeten kelen ketata eta majen hetero ten artan antideta eta eta h	(2177)	Gln422-Tyr435	
की विद्यालया के प्रकार कर के दिल्ला के किए के प्रकार के किए के प्रकार के किए के किए के किए के किए के किए किए क	(2201)	Arg426-Lys432	
अल्बल्या व्यवस्थात । स्वतं विद्या स्वतं । विद्या । विद्य	(2201)	Arg426-Gly431B	
द्रारा नेदेश कर महावस्था व्याप्त कर्मा विकास क्षेत्र प्राप्त कर विकास कर विकास कर विकास कर विकास कर विकास कर व	(2195)	Asn425-Lys432	
GCCCCTGGTGCACGGCCTGCTGGCCCTGATCTGGGACGA	(2201)	Consensus	
2241 2280			
१७ अववद्धारम् । दस्य १० ४० छ। १० ४५ छ। १० ४५ १० छ। १० ४५ १० छ।	(2229)	Ile424-Ala433	
क्ष्रेंद्रिक् <mark>रीतिकोम्प्रस्थितक</mark> ्ष्रियक्षेत्रक्षितिकार्	(2241)	Trp427-Gly431	
	(2217)	Gln422-Tyr435B	
नुकारवादावाद्यव्यवस्थात्वादाद्यावादावादावादावादावादावादावादावादावादावा	(2241)	Arg426-Gly431	
त्रेर प्रदायकार्यस्य प्रकारसम्बद्धाराम् सार्वास्य क्रायां स्वार्थकार्यक्षात्रस्य	(2223)	Ile423-Met434	
म्बल्लाहरूक्षात्रक्ष्माद्रक्ष्माद्रक्ष्मेद्रद्रत्ताम्बल्ला हर्ष्यक्षम् स्विन्नाद्वस्य विवर्धस्य	(2217)	Gln422-Tyr435	
व्यक्ताहर् अन्तिवाधिवाद्यं क्षेत्रमें संस्थित कालका कार्या कार्यक स्थान स्थान स्थान स्थान स्थान स्थान स्थान स	(2241)	Arg426-Lys432	
વૃદ્ધાં મહારા કામ લાકે કામ કામ માત્ર કામ માત્ર કામ માત્ર કામ માત્ર કામ	(2241)	Arg426-Gly431B Asn425-Lys432	
control of the contro	(2235) (2241)	Consensus	
CCTGCGCAGCCTGTGCCTGTTCAGCTACCACCGCCTGCGC	(2241)	consensus	
EIC AV			

FIG. 4K

		2281 232
Ile424-Ala433	(2269)	च गलाला हवा अमृत्विक संवा अस्ति अस्ति हो होते होते होते स्था स्था होते होते होते होते होते होते होते होते
Trp427-Gly431	(2281)	न्त्रहालस्यः अवक्रियक्षकान्त्रविद्यालयस्य विद्यान्त्रविद्यान्त्रविद्यान्त्रविद्यान्त्रविद्यान्त्रविद्यान्त्रवि
Gln422-Tyr435B	(2257)	स्तर्रद्रांद्रीयाच्यान्त्राद्रवार्याद्रोत्रयाद्रवार्याद्रवार्याद्रावाद्रवार्याद्रवार्याद्रवार्याद्रवार्याद्रवार
Arg426-Gly431	(2281)	वृत्यद्वित्तर्थकार्यकार्यकार्यकार्यकार्यद्वित्तर्थक्ष्यकार्यक्ष्यकार्यकार्यकार्यकार्यकार्यकार्यकार्यकार
Ile423-Met434	(2263)	वर्षयक्षांत्रम् स्टायात्रम् स्टायायका स्टायर भुद्रेत्वस्थरमस्याद्वेत्रस्य
Gln422-Tyr435	(2257)	च्याक्षरम् व्याप्तात्रभाषे कृति कृति । विद्यान स्थान । विद्या । विद्यान । विद्यान । विद्यान । विद्यान । विद्या
Arg426-Lys432	(2281)	૱૽ૣૼૡઌઌ૾ૡૹૡઌૺૹૡૹૡ૽ૡ૽ઌઌૹઌ૽૽૾૽ૢૡૡ૽૱ૡૡૡૡૡૡૡૡૡૡ
Arg426-Gly431B	(2281)	व्यक्षित्रकार में इंग्लेश विभाग विद्यवस्थ के अंतर के अस्ति स्वाक विद्यार स्वाक विद्यार
Asn425-Lys432	(2275)	व्यावकारतः १५५७ को रसः १४६० विद्यालया । याद्यान् १४८० । याद्यान् १४८० । या
Consensus	(2281)	GACCTGATCCTGATCGCCGCCCGCATCGTGGAGCTGCTGC
		2321 2360
Ile424-Ala433	(2309)	वर्षात्रद्वत्त्वत्त्रवर्षः वर्षात्रवर्षात्रवर्षात्रवर्षात्रवर्षात्रवर्षात्रवर्षात्रवर्षात्रवर्षाः
Trp427-Gly431	(2321)	बोर्स्चरात्रीन्त्विकः स्राम्स्यविकार्यक्ष्यायः स्राप्तायः स्राप्तायः स्राप्तायः
Gln422-Tyr435B	(2297)	सर्वा विदेशका सम्बद्धां संबंधित वासका का अवस्ति के अध्यति सम्बद्धां सम्बद्धाः
Arg426-Gly431	(2321)	सक्ष्याचीर विद्याराचन भूमानि भूमाना स्थापन अन्तर्भावक मानानि हो। स्थापन दिय
Ile423-Met434	(2303)	<u> बल्लाम् इत्याचनम्बद्धाः स्तर्भातः व्यामस्तर्भातः इत्यामस्तर्भातः । स्तर्भातः स्तर्भातः स्तर्भातः स्तर्भातः स</u>
Gln422-Tyr435	(2297)	माञ्ज्ञमाञ्चलकाराकारमञ्जलन्त्रात्।
Arg426-Lys432	(2321)	दार त्यावादीक विकास मानिकार अन्यदारावाक स्वतः अन्तिक अन्यविकास स्वतः विकास
Arg426-Gly431B	(2321)	्रविष्युव्यविष्युवर्ष्युक्तात्रविष्युत्रिय्वविष्युक्तात्रिय्वय्युक्तात्रिय्वय्युक्तात्रिय्वय्युक्तात्रियः
Asn425-Lys432	(2315)	प्रवित्याचार्वत्यामुद्राच्याव्याचार्याच्याच्यात्र्याच्यात्र्यात्र्यात्र्यात्र्यात्र्यात्र्यात्र्यात्र्यात्र्या
Consensus	(2321)	GCCGCCGCGCTGGGAGGCCCTGAAGTACTGGGGCAACCT
0000545	(2321)	2361 2400
Ile424-Ala433	(2349)	acticated and any southern and a second seco
Trp427-Gly431	(2361)	etepestatus que meter en el suferon en hacinatologiam accidente en
Gln422-Tyr435B	(2337)	description as released the property of the experience of the same
Arg426-Gly431	(2361)	the facilitation of anythologist facilities and when the property continues to
Ile423-Met434	(2343)	estine outenation defendation of the meaning of the contraction of the
Gln422-Tyr435	(2337)	etaindesparationalisticamination and introductional electric
Arg426-Lys432	(2361)	assignmental elementelementelementelementelementelementelementelemente
Arg426-Gly431B	(2361)	The state of the s
Asn425-Lys432	(2355)	nastrialistiche delicitation dem dem delicitation delicit
Consensus	(2361)	de seconomica del construir de seconomica de
Consensus	(2361)	GCTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTG 2401 2440
Ile424-Ala433	(2200)	
Trp427-Gly431	(2389)	in the second of the contraction is the second seco
	(2401)	ांद्राराक्षात्रकात्रकात्रक्षात्रवात्रकात्रात्रकात्रकात्रकात्रकात्रकात्रका
Gln422-Tyr435B	(2377)	jedi. komananyangananan aminakan panan adam.
Arg426-Gly431	(2401)	ादाहर प्रवासिद्धार्यद्वार विकासिक्ता । विकास विवास स्थापन करान्य । विकास स्थापन करान्य । विकास स्थापन करान्य स
Ile423-Met434	(2383)	ावर्गः ।क्षान्यविद्योष्ट्राक्षाः । भग्नान्य द्रव्यक्षाः अर्थे व्यवसम्बद्धाः स्तित्वर ।
Gln422-Tyr435	(2377)	भवन्तर्यकार्यकार्यक्षेत्रस्य विश्वसंभित्रस्य स्वतास्य स्वतास्य स्वतास्य स्वतास्य स्वतास्य स्वतास्य स्वतास्य स्
Arg426-Lys432	(2401)	પાસારા તેમ ત્યાર વર્ષ ા કરાયા છે. તેમ
Arg426-Gly431B	(2401)	notice englestations in the second continues and con
Asn425-Lys432		भवतः अक्षामिविक्षम् वर्गातम् । वर्गादिः अस्य अस्य अस्य वर्गानस्य । अस्य अस्य अस्य ।
Consensus	(2401)	AGCCTGTTCGACGCCATCGCCATCGCCGTGGCCGAGGGCA
T1-404 P1 400		2441 2480
Ile424-Ala433		भ्यत्। १७ वर्षः । ११ १८ वर्षः ११ १९ वर्षः । १९ वर्षः १९ वर्षः । १९ वर्षः १९ वर्षः १९ वर्षः १९ वर्षः १९ वर्षः १
Trp427-Gly431		न्यांकाराम्बद्धानम्भक्तम् विभिन्नकारमञ्ज्ञानम् । सम्बद्धानम् अस्य अस्य वि
Gln422-Tyr435B		<i>ૺૡ૽૽૽૱ૡ૽૽ૡ૽૽ઌ૽ઌ૽ૡ૽ૺઌ૽૽ૡ૽ૡઌૡ૽ૡૡઌઌઌઌ૽ૡ૽ઌ૽</i> ઌઌૡૡૡૡૡૡૡૡૡૡ
Arg426-Gly431		्राहर । अवस्थित क्षेत्र क्षेत्र इत्र क्षेत्र क
Ile423-Met434		्ववास्थ्रव र्वत्यं ४३व्यापटिक्रां वर्त्या ८०० व्यावस्थातस्य स्वतं वर्षात्रस्य
Gln422-Tyr435		त्यवात्रकाताम् प्रकृतिक अवस्थित । व्यवस्थित । स्वतः व्यवस्थान
Arg426-Lys432		लेलका/अवस्था:प्रमुख्यक्षलका/वास्त्रविकार्यका/अवस्थान्यक्ष्यका/अवस्थान्यक्ष्यका
Arg426-Gly431B	(2441)	ाः व्याप्तराचनः विशेषाः चेत्रप्रदान्यन् । जन्मेन अस्ति । जन्मेन । जन्मेन । जन्मेन । जन्मेन । जन्मेन । जन्मेन ।
Asn425-Lys432	(2435)	द्रत्या/तृत्वत्र्वाप्रदेशका देशकृतिके प्रविद्यात्रा कृतिवास कार्यस्य द्रात्वा विद्या ।
Consensus		CCGACCGCATCATCGAGGTGGCCCAGCGCATCGGCCGCGC
		2481 2520
Ile424-Ala433	(2469)	व्याप्तित्वेपत्रात्रात्रकारम् । त्राप्तात्रात्रात्रात्रात्रात्रात्रात्रात्रात
		FIG. AI

FIG. 4L

Trp427-Gly431	(2481)	त्रमारचीत्मर स्वे प्रकृतिकारमञ्जलेखात् । स्वयं विद्याने स्वयं स्वयं विद्याने । स्वयं विद्याने स्वयं विद्याने ।
Gln422-Tyr435B	(2457)	क्षेत्राञ्चरभवत्त्राक्ष्मप्रविद्धवात्त्रावद्यप्रभवविद्यक्षमण्यविद्यम्भकत्त्रात
Arg426-Gly431	(2481)	क्षा विक्रम् १८६० १८८ भूग व र्षा द्वारा द्वारा द्वारा द्वारा द्वारा द्वारा होता है।
Ile423-Met434	(2463)	लेक्ष्रेंबाक्षरम्भाग्याक्षेत्रकार्यस्थानार्वस्थात्रात्रात्रात्रात्रात्रात्रात्रात्रात्र
Gln422-Tyr435	(2457)	रुक्तांदरी <u>लाइन्दर्क</u> ार्यकृतिस्त्रीत्राहरू संदर्शाद्वादात्वर साम्बन्धरू साम्बन्धर स्वतास्त्रीत्वर साम्बन्धर स्वतास
Arg426-Lys432	(2481)	इत्यादिक्यं वृद्धां त्राप्त्रं वृद्धां विकारित विकारित विकारित विकारित विकारित विकारित विकारित विकारित विकारित
Arg426-Gly431B	(2481)	समाहित्सावत्। १% १४ वर्षाच्या व्यवस्थितः १५ १५ १५ १५ १५ १५ १५ १५ १५ १५ १५ १५ १५
Asn425-Lys432	(2475)	के माहित्यम् स्वतिक के हिन्द्र सिट्टिट्टिट्टिट्टिट्टिट्टिट्टिट्टिट्टिट्
Consensus	(2481)	CTTCCTGCACATCCCCCGCCGCATCCGCCAGGGCTTCGAG
		2521 2541
Ile424-Ala433	(2509)	संबद्धारिकार्वाद में बेहित में हैं में भी का अपने हैं कि अपने
Trp427-Gly431	(2521)	Relationers designation of the
Gln422-Tyr435B	(2497)	अवदार्शनंतरम् प्रतिकेतिन्त्रम् । स्ट्रान्तरम् ।
Arg426-Gly431	(2521)	अमेले हार है में में हुए हैं है
Ile423-Met434	(2503)	ल्लिक्ट्रिक्
Gln422-Tyr435	(2497)	signed and extraction sector services in the sector and the sector secto
Arg426-Lys432	(2521)	Eddicate Acid (Edition Acids Total
Arg426-Gly431B	(2521)	Seineles (cydnomie) co Valenticiente
Asn425-Lys432	(2515)	ACIENCATION OF SECTIONAL SECTION AND SECTI
Consensus	(2521)	CGCGCCCTGCTGTAACTCGAG

FIG. 4M

WO 00/39303	28	/ ₁ 65 PCT/US99/31272
Leu122-Ser199-Tryp427-Gly431	(3)	GAATTCCCCACCATGGATGCAATGAAGAGA
Val127-Asn195-Arg426-Gly431	(1)	a transfer that we have a first and the second seco
Val120-Thr202-lle424-Ala433	(1)	the second section of the second section is a second section of the second section of the second section is a second section of the section of
Leu122-Ser199-Arg426-Lys432	(1)	GAATTCGCCACCATGGATGCAATGAAGAGA
Leu122-Ser199-Arg426-Gly431	(1)	GAATTCGCCACCATGGATGCAATGAAGAGA
Lys121-Val200-Asn425-Lys432	(1)	and the state of t
Val120-Ile201-Ile424-Ala433	(1)	And the second for the second
		a consideration of the constant of the constan
Vall20-Ile201B-Ile424-Ala433	(1)	
Consensus	(1)	GAATTCGCCACCATGGATGCAATGAAGAGA 31 60
Leu122-Ser199-Tryp427-Gly431	(31)	GGGCTCTGCTGTGTGCTGCTGTGTGGA
Val127-Asn195-Arg426-Gly431	(31)	GGGCTCTGCTGTGTGCTGCTGTGTGGA
Val120-Thr202-Ile424-Ala433	(31)	GGGCTCTGCTGTGTGCTGCTGTGTGGA
Leu122-Ser199-Arg426-Lys432	(31)	GGGCTCTGCTGTGTGCTGCTGTGTGGA
Leu122-Ser199-Arg426-Gly431	(31)	GGGCTCTGCTGTGTGCTGCTGTGTGGA
Lys121-Va1200-Asn425-Lys432	(31)	GGGCTCTGCTGTGTGCTGCTGTGGA
Val120-Ile201-Ile424-Ala433	(31)	GGGCTCTGCTGTGTGCTGCTGTGTGGA
Val120-Ile201B-Ile424-Ala433	(31)	GGGCTCTGCTGTGTGCTGCTGTGTGGA
Consensus	(31)	GGGCTCTGCTGTGCTGCTGTGTGGA
•		61 90
Leu122-Ser199-Tryp427-Gly431	(61)	GCAGTCTTCGTTTCGCCCAGCGCCGTGGAG
Val127-Asn195-Arg426-Gly431	(61)	GCAGTCTTCGTTTCGCCCAGCGCCGTGGAG
Val120-Thr202-Ile424-Ala433	(61)	GCAGTETTCGTTTCGCCCAGCGCCGTGGAG
Leu122-Ser199-Arg426-Lys432	(61)	GCAGTCTTCGTTTCGCCCAGCGCCGTGGAG
Leu122-Ser199-Arg426-Gly431	(61)	GCAGTCTTCGTTTCGCCCAGCGCCGTGGAG
Lys121-Val200-Asn425-Lys432	(61)	GCAGTCTTCGTTTCGCCCAGCGCGTGGAG
Val120-Ile201-Ile424-Ala433	(61)	GCAGTCTTCGTTTCGCCCAGCGCCGTGGAG
Val120-Ile201B-Ile424-Ala433	(61)	GCAGTCTTCGTTTCGCCCAGCGCCGTGGAG
Consensus	(61)	GCAGTCTTCGTTTCGCCCAGCGCCGTGGAG
		91 120
Leu122-Ser199-Tryp427-Gly431	(91)	AAGCTGTGGGTGACCGTGTACTACGGCGTG
Val127-Asn195-Arg426-Gly431	(91)	AAGCTGTGGGTGACCGTGTACTACGGCGTG
Val120-Thr202-Ile424-Ala433	(91)	AAGCTGTGGGTGACCGTGTACTACGGCGTG
Leu122-Ser199-Arg426-Lys432	(91)	AAGCTGTGGGTGACCGTGTACTACGGCGTG
Leu122-Ser199-Arg426-Gly431	(91)	The second secon
Lys121-Va1200-Asn425-Lys432	(91)	AAGCTETGGGTGACCGTGTACTACGGCGTG
Val120-Ile201-Ile424-Ala433	(91)	AAGGTGTGGGTGACCGTGTACTACGGCGTG
Val120-Ile201B-Ile424-Ala433	(91)	AAGCTGTGGGTGACCGTGTACTACGGCGTG
Consensus	(91)	AAGCTGTGGGTGACCGTGTACTACGGCGTG
00.00 \$1.15 4.5	(,	121 150
Leu122-Ser199-Tryp427-Gly431	(121)	CCCGTGTGGAAGGAGGCCACCACCCTG
Val127-Asn195-Arg426-Gly431	(121)	CCCGTGTGGAAGGAGGCCACCACCACCCTG
Val120-Thr202-Ile424-Ala433		CCCGTGTGGAAGGAGGCCACCACCACCTG
Leu122-Ser199-Arg426-Lys432		CCCGTGTGCAAGGAGGCCACCACCACGCTG
Leu122-Ser199-Arg426-Gly431		CCCTCTGGAAGGAGGCCACCACCACCCTG
Lys121-Va1200-Asn425-Lys432		CECGIGIGGAGGAGGCCACCACCACCACCACCACCACCACCACCACC
Val120-Ile201-Ile424-Ala433	(121)	CCCGTGTGGAAGGAGGCCACCACCACCCTG
Val120-Ile201B-Ile424-Ala433	(121)	CCCGTGTGGAAGGAGGCCACCACCACCCTG
Consensus		CCCGTGTGGAAGGAGGCCACCACCACCCTG
•		151 180
Leu122-Ser199-Tryp427-Gly431	(151)	TTCTGCGCCAGCGACGCCAAGGCCTACGAC
Val127-Asn195-Arg426-Gly431	(151)	TTCTGCGCCAGCGACGCCAAGGCCTACGAC
Val120-Thr202-Ile424-Ala433	(151)	TTCTGCGCCAGCGACGCCTACGAC
Leu122-Ser199-Arg426-Lys432	(151)	TTCTGEGCCAGCGACGCCAAGGCCTACGAC
Leul22-Ser199-Arg426-Gly431	(151)	TTCTGCGCCAGCGACGCCAAGGCCTACGAC
Lys121-Va1200-Asn425-Lys432	(151)	TTCTGCGCCAGCGACGCCAAGGCCTACGAC

WO 00/20202		/ 65 PCT/US99/31272
WO 00/39303	29	7 03
Val120-Ile201-Ile424-Ala433	(151)	
Vall20-Ile201B-Ile424-Ala433	(151)	TTCTGCGCCAGCGACGCCAAGGCCTACGAC
Consensus	(151)	TTCTGCGCCAGCGACGCCAAGGCCTACGAC 181 210
Leu122-Ser199-Tryp427-Gly431	(181)	
Vall27-Asn195-Arg426-Gly431	(181)	ACCGAGGTGCACAACGTGTGGGCCACCCAC
Val120-Thr202-Ile424-Ala433	(181)	and the state of t
Leu122-Ser199-Arg426-Lys432	(181)	ACCGAGGTGCACAACGTGTGGGCCACCCAC
Leu122-Ser199-Arg426-Gly431	(181)	ACCGAGGTGCACACGTGTGGGCCACCCAC
Lys121-Val200-Asn425-Lys432	(181)	ACCGAGGTGCACAACGTGTGGGCCACCCAC
Val120-Ile201-Ile424-Ala433	(181)	ACCGAGGTGCACAACGTGTGGGCCACCCAC
Val120-Ile201B-Ile424-Ala433	(181)	
Consensus	(181)	ACCGAGGTGCACAACGTGTGGGCCACCCAC
Taval 202 Cara 200 May 407 C1 423		211 240
Leu122-Ser199-Tryp427-Gly431		GCCTGCGTGCCCACCGACCCCAACCCCCAG
Val127-Asn195-Arg426-Gly431 Val120-Thr202-Ile424-Ala433	(211)	GCCTGCGTGCCACCGACCCCAACCCCGAG GCCTGGGTGCCACCGACCCCAACCCCAAG
Leu122-Ser199-Arg426-Lys432		GCCTGCGTGCCCACCCCAACCCCCAG
Leu122-Ser199-Arg426-Gly431	(211)	GCCTGCGTGCCCACCCCAACCCCCAG
Lys121-Val200-Asn425-Lys432	(211)	the contract of the contract o
Val120-Ile201-Ile424-Ala433	(211)	GCCTGCGTGCCCACCGACCCCCAG
Val120-Ile201B-Ile424-Ala433	(211)	GCCTGCGTGCCCACCGACCCCAG
Consensus	(211)	GCCTGCGTGCCCACCGACCCCAACCCCCAG
		241 270
Leu122-Ser199-Tryp427-Gly431	(241)	the Court of the C
Val127-Asn195-Arg426-Gly431	(241)	GAGATEGTGCTGGAGAACGTGACCGAGAAC
Val120-Thr202-Ile424-Ala433	(241)	GAGATCGTGCTGGAGAACGTGACCGAGAAC
Leu122-Ser199-Arg426-Lys432	(241)	GAGATEGTGCTGGAGAACGTGACCGAGAAC GAGATCGTGCTGGAGAACGTGACCGAGAAC
Leu122-Ser199-Arg426-Gly431 Lys121-Va1200-Asn425-Lys432	(241) (241)	GAGATEGTGETGGAGAACGTGACCGAGAAC
Val120-Ile201-Ile424-Ala433	(241)	GAGAICGIGCTGGAGAACGIGACCGAGAAC
Val120-Ile201B-Ile424-Ala433	(241)	GAGATCGTGCTGGAGAACGTGACCGAGAAC
Consensus		GAGATCGTGCTGGAGAACGTGACCGAGAAC
	•	271 300
Leu122-Ser199-Tryp427-Gly431	(271)	TTCAACATGTGGAAGAACAACATGGTGGAG
Val127-Asn195-Arg426-Gly431	(271)	TTCAACATGTGGAAGAACAACATGGTGGAG
Val120-Thr202-Ile424~Ala433	(271)	TTCAACATGTGGAAGAACAACATGGTGGAG
Leu122-Ser199-Arg426-Lys432	(271)	TTCAACATGTGGAAGAACAACATGGTGGAG
Leu122-Ser199-Arg426-Gly431	(271)	TTCAACATGTGGAAGAACAACATGGTGGAG
Lys121-Val200-Asn425-Lys432	(271)	TTCAACATGTGGAAGAACAACATGGTGGAG
Val120-Ile201-Ile424-Ala433 Val120-Ile201B-Ile424-Ala433	(271)	TTCAACATGTGGAAGAACAACATGGTGGAG TTCAACATGTGGAAGAACAACATGGTGGAG
Consensus	(271) (271)	
Consensus	(2/1)	301 330
Leu122-Ser199-Tryp427-Gly431	(301)	
Val127-Asn195-Arg426-Gly431	(301)	\$250 of principle of the committee of th
Val120-Thr202-Ile424-Ala433		CAGATECACGAGGACATCATCAGCCTGTGG
Leu122-Ser199-Arg426-Lys432	(301)	CAGATGCACGAGGACATCATCAGCCTGTGG
Leu122-Ser199-Arg426-Gly431		CAGATGUACGAGGACATCATCAGCCTGTGG
Lys121-Val200-Asn425-Lys432		CAGATGCACGAGGACATCATCAGCCTGTGG
Val120-Ile201-Ile424-Ala433		CAGATGCACGAGGACATCATCAGCCTGTGG
Val120-Ile201B-Ile424-Ala433	(301)	
Consensus	(301)	CAGATGCACGAGGACATCATCAGCCTGTGG
Leu122-Ser199-Tryp427-Gly431	/331:	331 360 SACCAGAGCCTGAAGCCCTGCGTGAAGCTG
Vall27-Asn195-Arg426-Gly431		GACCAGAGCCTGAAGCCCGGAGCCTGGAGCCTGGAGCCTGAAGCCTGAAGCCTGAAGCCTGAAGCCTGAAGCTG
Val120-Thr202-Ile424-Ala433		GACCAGAGCCTGAAGCCCTGCGTG
	, 1	ವರ್ಣ ನಿರ್ವಹಿಸುವ ಪ್ರವಾಧ ಪ್ರವಾಧವನ್ನು ಪ್ರವಾಧ ಪ್ರವಹಿಸುತ್ತವೆ. ಪ್ರ

WO 00/39303	30	/ 65 PCT/US99/31272
Leu122-Ser199-Arg426-Lys432	(331)	GACCAGAGCCTGAAGCCCTGCGTGAAGCTG
Leu122-Ser199-Arg426-Gly431	(331)	GACCAGAGCCTGAAGCCCTGCGTGAAGCTG
Lys121-Val200-Asn425-Lys432	(331)	GACCAGAGCCTGAAGCCCTGCGTGAA
Val120-Ile201-Ile424-Ala433	(331)	GACCAGAGCCTGAAGCCCTGCGTG
Val120-Ile201B-Ile424-Ala433	(331)	GACCAGAGCCTGAAGCCCTGCGTG
•	(331)	GACCAGAGCCTGAAGCCCTGCGTGAAGCTG
Consensus	(331)	
	12611	361 390
Leu122-Ser199-Tryp427-Gly431	(361)	GG
Val127-Asn195-Arg426-Gly431	(361)	ACCCCCTGTGCGTGGGGCAGGGAACTGC
Val120-Thr202-Ile424-Ala433	(355)	GG
Leu122-Ser199-Arg426-Lys432	(361)	GG
Leu122-Ser199-Arg426-Gly431	(361)	<u>G</u> G
Lys121-Val200-Asn425-Lys432	(357)	GG
Val120-Ile201-Ile424-Ala433	(355)	
Val120-Ile201B-Ile424-Ala433	(355)	
Consensus	(361)	GG
	•	391 420
Leu122-Ser199-Tryp427-Gly431	(363)	CAACAGCGTGATCACCCAGGCCTGCCCC
Val127-Asn195-Arg426-Gly431	(391)	AACACCAGCGTGATCACCCAGGCCTGCCCC
Val120-Thr202-Ile424-Ala433	(357)	CGCCGCCACCCAGGCCTGCCCC
Leu122-Ser199-Arg426-Lys432	(363)	CAACAGCGTGATCACCCAGGCCTGCCCC
Leu122-Ser199-Arg426-Gly431	(363)	CAACAGCGTGATCACCCAGGCCTGCCCC
Lys121-Val200-Asn425-Lys432	(359)	CCCCEGTGATCACCCAGGCCTGCCCC
Val120-Ile201-Ile424-Ala433	(355)	GECGCATCACCCAGGCCTGCCCC
Vall20-Ile201B-Ile424-Ala433	(355)	CCEGGCATCACCCAGGCCTGCCCC
Consensus	(391)	CA CAGCGTGATCACCCAGGCCTGCCCC
•		421 450
Leu122-Ser199-Tryp427-Gly431	(391)	AAGGTGAGCTTCGAGCCCATCCCATCCAC
Val127-Asn195-Arg426-Gly431	(421)	AAGGTGAGCTTCGAGCCCATCCCCATCCAC
Val120-Thr202-Ile424-Ala433	(379)	AAGGTGAGCTTCGAGCCCATCCCATCCAC
Leu122-Ser199-Arg426-Lys432	(391)	AAGGTGAGCTTCGAGCCCATCCCCATCCAC
Leu122-Ser199-Arg426-Gly431	(391)	AAGGTGAGCTTEGAGCCCATCCCCATCCAC
Lys121-Val200-Asn425-Lys432	(385)	AAGGTGAGCTTCGAGCCCATCCCAC
Val120-Ile201-Ile424-Ala433	(379)	AAGGUGAGO WAGAGCCGATICCAC
Val120-Ile201B-Ile424-Ala433	(379)	AAGGTGAGCTTCGAGCCCATCCCATCCAC
Consensus	(421)	AAGGTGAGCTTCGAGCCCATCCCCATCCAC
consensus	(321)	451 480
I au 122 Ca = 100 T = 427 Cl 421	(421)	TACTEGEROCCECCECCECTCCCATCCTC
Leu122-Ser199-Tryp427-Gly431	(421)	
Val127-Asn195-Arg426-Gly431	(451)	TACIGOGOCOCCGCCGCCTTCGCCATCCTG
Val120-Thr202-Ile424-Ala433	(409)	TACTECGCCGCCGGCTTCGCCATCCTG
Leu122-Ser199-Arg426-Lys432	(421)	TACTECCECCEGGCGGCTTCGCCATCCTG
Leu122-Ser199-Arg426-Gly431	(421)	TACTGEGEGEGGETTEGCCATCETG
Lys121-Va1200-Asn425-Lys432	(415)	TACTICOGICECECCGGCTTCGCCATCCTG
Val120-Ile201-Ile424-Ala433	(409)	WARTICES CONSECCED CONTESTS
Vall20-Ile201B-Ile424-Ala433	(409)	TACTCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
Consensus	(451)	TACTGCGCCCCGCCGGCTTCGCCATCCTG
		481 510
Leu122-Ser199-Tryp427-Gly431	(451)	AAGTGCAACGACAAGAAGTTCAACGGCAGC
Val127-Asn195-Arg426-Gly431	(481)	AAGTGCAACGACAAGAAGTTCAACGGCAGC
Val120-Thr202-Ile424-Ala433		AAGTGCAAGGACAAGAAGTTCAACGGCAGC
Leu122-Ser199-Arg426-Lys432		AAGTGCAACGACAAGAAGTTCAACGGCAGC
Leu122-Ser199-Arg426-Gly431		AAGTGCAACGACAAGAAGTTCAACGGCAGC
Lys121-Val200-Asn425-Lys432		AAGTECAACAAGAAGTTCAACGGCAGC
Val120-Ile201-Ile424-Ala433		AAGTGCAACGACAAGAAGTTCAACGGCAGC
Val120-11e201-11e424-A1a433 Val120-11e201B-11e424-A1a433		
	(439)	AAGTGCAACGACAAGAAGTTCAACGGCAGC
Consensus	(481)	AAGTGCAACGACAAGAAGTTCAACGGCAGC
		511 540

		Demilion #1973
WO 00/39303	.31	/ 65 PCT/US99/31272
Leu122-Ser199-Tryp427-Gly431	(481)	GGCCCTGCACCAACGTGAGCACCGTGCAG
Val127-Asn195-Arg426-Gly431	(511)	GGCCCTGCACCAACGTGAGCACCGTGCAG
Val120-Thr202-Ile424-Ala433	(469)	GGCCCTGCACCAACGTGAGCACCGTGCAG
Leu122-Ser199-Arg426-Lys432	(481)	GGCCCCTGCACCAACGTGAGCACCGTGCAG
Leu122-Ser199-Arg426-Gly431	(481)	GGCCCTGCACCAACGTGAGCACCGTGCAG
Lys121-Val200-Asn425-Lys432	(475)	GGCCCCTGCACCAACGTGAGCACCGTGCAG
Val120-Ile201-Ile424-Ala433	(469)	GGCCCTGCACCAACGTGAGCACCGTGCAG
Val120-Ile201B-Ile424-Ala433	(469)	GGCCCTGCACCAACGTGAGCACCGTGCAG
Consensus	(511)	GGCCCCTGCACCAACGTGAGCACCGTGCAG
00110011040	(322)	541 570
Leu122-Ser199-Tryp427-Gly431	(511)	TGCACCCACGCCATCCGCCCCGTGGTGAGC
Val127-Asn195-Arg426-Gly431	(541)	TGCACCCACGGCATCCGCCCCGTGGTGAGC
Val120-Thr202-Ile424-Ala433	(499)	TECACCCACGGCATCCGCCCCGTGGTGAGC
		TGCACCCACGCCATCCGCCCCGTGGTGAGC
Leu122-Ser199-Arg426-Lys432	(511)	- And Strategies and Andreas and the Control of the
Leu122-Ser199-Arg426-Gly431	(511)	TGCACCCACGGCATCCGCCCCGTGGTGAGC
Lys121-Val200-Asn425-Lys432	(505)	TGCACCCACGGCATCCGCCCCGTGGTGAGC
Val120-Ile201-Ile424-Ala433	(499)	TGCACCCACGGCATCCGCCCCGTGGTGAGC
Val120-Ile201B-Ile424-Ala433	(499)	TGCACCCACGGCATCCGCCCCGTGGTGAGC
Consensus	(541)	TGCACCCACGGCATCCGCCCCGTGGTGAGC
		571 600
Leu122-Ser199-Tryp427-Gly431	(541)	
Val127-Asn195-Arg426-Gly431	(571)	ACCCAGCTGCTGCTGAACGCCAGCCTGGCC
Val120-Thr202-Ile424-Ala433	(529)	ACCCAGCTGGTGCTGAACGGCAGCCTGGCC
Leu122-Ser199-Arg426-Lys432	(541)	ACCCAGCTGCTGCAACGCCAGCCTGGCC
Leu122-Ser199-Arg426-Gly431	(541)	ACCCAGCTGCTGCTGAACGGCAGCCTGGCC
Lys121-Val200-Asn425-Lys432	(535)	Transferrence and the second s
Val120-Ile201-Ile424-Ala433	(529)	
Vall20-Ile201B-Ile424-Ala433	(529)	ACCOAGCTGCTGCTGAACGGCAGCCTGGCC
Consensus	(571)	- before the second to the sec
00.130.1343	(3,1)	601 630
Leu122-Ser199-Tryp427-Gly431	(571)	GAGGAGGCGTGGTGATCCGCAGCGAGAAC
Val127-Asn195-Arg426-Gly431	(601)	GAGGAGGGCGTGGTGATCCGCAGCGAGAAC
Val120-Thr202-Ile424-Ala433	(559)	
	(571)	GAGGAGGGCGTGGTGATCCGCAGCAGCAC
Leu122-Ser199-Arg426-Lys432	(571)	
Leu122-Ser199-Arg426-Gly431		GAGGAGGGCTGGTGATCCGCGAGAAC
Lys121-Val200-Asn425-Lys432	(565)	
Val120-Ile201-Ile424-Ala433	(559)	GAGGAGGCGTGGTGATCCGCAGCGAGAAC
Vall20-Ile201B-Ile424-Ala433	(559)	GAGGAGGEGAGGTGATCCGCAGCGAGAAC
Consensus	(601)	
		631 660
Leu122-Ser199-Tryp427-Gly431	(601)	TTCACCGACAACGCCAAGACCATCATCGTG
Val127-Asn195-Arg426-Gly431	(631)	TTCACCGACAACGCCAAGACCATCATCGTG
Val120-Thr202-Ile424-Ala433	(589)	TTCACCGACAACGCCAAGACCATCATCGTG
Leu122-Ser199-Arg426-Lys432	(601)	TTCACCGACAACGCCAAGACCATCATCGTG
Leu122-Ser199-Arg426-Gly431	(601)	TTCACCGACAACGCCAAGACCATGATGGTG
Lys121-Val200-Asn425-Lys432	(595)	TTEAGGGACAACGGGAAGACCATGATGGTG
Val120-Ile201-Ile424-Ala433	(589)	TTCACCGACAACGCCAAGACCATCATEGEG
Val120-Ile201B-Ile424-Ala433	(589)	TTCACCGACAACGCCAAGACCATCATCGTG
Consensus	(631)	TTCACCGACAACGCCAAGACCATCATCGTG
555564	/	661 690
Leul22-Ser199-Tryp427-Gly431	(631)	CAGCTGAAGGAGAGCGTGGAGATCAACTGC
Vall27-Asn195-Arg426-Gly431		CAGCTGAAGGAGAGCGTGGAGATCAACTGC
Val120-Thr202-Ile424-Ala433		CAGCTGAAGGAGGCGTGGAGATCAACTGC
Leu122-Ser199-Arg426-Lys432		CAGCTGAAGGAGCGTGGAGATCAACTGC
-		CAGCTGAAGGAGAGCGTGGAGATCAACTGC CAGCTGAAGGAGAGCGTGGAGATCAACTGC
Leu122-Ser199-Arg426-Gly431	(631)	
Lys121-Va1200-Asn425-Lys432 Va1120-Ile201-Ile424-Ala433	(625)	CAGCTGAAGGAGAGCGTGGAGATCAACTGC
vai120-116201-116424-A18433	(019)	CAGCIGAAGGAGAGCGTGGAGATCAACIGC

WO 00/39303	32	/ 65	PCT/US99/31272
val120-11e2018-11e424-Ala433	(619)	CAGCTGAAGGAGAGCGTG	GAGATCAACTGC
Consensus	(661)	CAGCTGAAGGAGAGCGTG	
	, ,	691	720
Leu122-Ser199-Tryp427-Gly431	(661)	ACCCGCCCCAACAACAAC	
Val127-Asn195-Arg426-Gly431	(691)	ACCCCCCCAACAACAAC	
Val120-Thr202-Ile424-Ala433	(649)	ACCCGCCCAACAACAACA	
Leu122-Ser199-Arg426-Lys432	(661)	ACCCCCCCAACAACAACA	
Leu122-Ser199-Arg426-Gly431	(661)	ACCCGCCCCAACAACAAC	
Lys121-Val200-Asn425-Lys432	(655)	ACCEGECCCAACAACAAC	
Val120-Ile201-Ile424-Ala433	(649)	ACCCCCCCAACAACAACA	
Vall20 11e201 11e424 Ala433 Vall20-Ile201B-Ile424-Ala433	(649)	ACCCGCCCCAACAACAACA	
Consensus	(691)	ACCCGCCCCAACAACAACA	
Consensus	(0)1)	721	750
I 0122 Co.w.100-Marria 427-Clar421	(691)		
Leu122-Ser199-Tryp427-Gly431			
Val127-Asn195-Arg426-Gly431	(721)	ATCACCATCGGCCCCGGCC	
Vall20-Thr202-Ile424-Ala433	(679) (691)	ATCACCATCGGCCCCGGCC	
Leu122-Ser199-Arg426-Lys432		ATCACGATCGGCCCCGGCC	
Leu122-Ser199-Arg426-Gly431	(691)	ATCACCATCGGCCCCGGCC	
Lys121-Val200-Asn425-Lys432	(685)	ATCACCATCGGCCCCGGCC	
Val120-Ile201-Ile424-Ala433	(679)	ATCACCATCGGCCCCGGCC	
Vall20-Ile201B-Ile424-Ala433	(679)	ATCACCATCGGCCCCGGCC	
Consensus	(721)	ATCACCATCGGCCCCGGCC	· · · · · · · · · · · · · · · · · · ·
		751	780
Leu122-Ser199-Tryp427-Gly431	(721)	GCCACCGGCGACATCATCG	
Val127-Asn195-Arg426-Gly431	(751)	GCCACCGGCGACATCATCG	
Val120-Thr202-Ile424-Ala433	(709)	GCCACCGGCGACATCATCC	
Leu122-Ser199-Arg426-Lys432	(721)	GCCACCGGCGACATCATCG	
Leu122-Ser199-Arg426-Gly431	(721)	GCCACCGGCGACATCATCG	
Lys121-Val200-Asn425-Lys432	(715)	GCCACCGGCGACATCATCG	The state of the s
Val120-Ile201-Ile424-Ala433	(709)	GCCACEGGCGACATCATCG	The state of the s
Vall20-Ile201B-Ile424-Ala433	(709)	GCCACCGGCGACATCATCG	
Consensus	(751)		
		781	810
Leu122-Ser199-Tryp427-Gly431	(751)	CAGGCCCACTGCAACATCA	
Val127-Asn195-Arg426-Gly431	(781)	CAGGCCCACTGCAACATCA	
Val120-Thr202-Ile424-Ala433	(739)	CAGGCCCACTGCAACATCA	
Leu122-Ser199-Arg426-Lys432	(751)	CAGGECCACTGCAACATCA	
Leu122-Ser199-Arg426-Gly431	(751)	CAGGCCCACTGCAACATCA	GCGGCGAGAAG
Lys121-Val200-Asn425-Lys432	(745)	CAGGCCCACTGCAACATCA	GCGGCGAGAAG
Val120-Ile201-Ile424-Ala433	(739)	CAGGCCCACTGCAACATCA	GCGGCGAGAAG
Val120-Ile201B-Ile424-Ala433	(739)	CAGGCCCACTGCAACATCA	GCGGCGAGAAG
Consensus	(781)	CAGGCCCACTGCAACATCA	GCGGCGAGAAG
		811	840
Leu122-Ser199-Tryp427-Gly431	(781)	TGGAACAACACCCTGAAGC	AGATCGTGACC
Val127-Asn195-Arg426-Gly431	(811)	TGGAACAACACCCTGAAGC	
Val120-Thr202-Ile424-Ala433	(769)	TGGAACAACACCCTGAAGC	AGATCGTGACC
Leu122-Ser199-Arg426-Lys432	(781)	TGGAACAACACCCTGAAGC	
Leu122-Ser199-Arg426-Gly431	(781)	TGGAACAACACCCTGAAGC	
Lys121-Val200-Asn425-Lys432	(775)	TGGAACAACACCCTGAAGC	
Val120-Ile201-Ile424-Ala433	(769)	TGGAACAACACCCTGAAGC	
Val120-Ile201B-Ile424-Ala433	(769)	TGGAACAACACCCTGAAGC	
Consensus	(811)	TGGAACAACACCCTGAAGC	
3353545	, /	841	870
Leu122-Ser199-Tryp427-Gly431	(811)	AAGCTGCAGGCCCAGTTCG	- ·
Val127-Asn195-Arg426-Gly431	(841)	AAGCTGCAGGCCCAGTTCG	
Vall20-Thr202-Ile424-Ala433	(799)	AAGCTGCAGGCCCAGTTCG	
Leu122-Ser199-Arg426-Lys432	(811)	AAGCTGCAGGCCCAGTTCG	
Double Collins Migazo-Dysasz	(011)	THOU TOUR DOUGLE LONG TICO	acuriting the

WO 00/39303 PCT/US99/31272 33 65 Leu122-Ser199-Arg426-Gly431 (811) AAGCTGCAGGCCCAGTTCGGCAACAAGACC (805) AAGCTGCAGGCCCAGTTCGGCAACAAGACC Lys121-Val200-Asn425-Lys432 (799) AAGCTGCAGGCCCAGTTCGGCAACAAGACC Val120-Ile201-Ile424-Ala433 Val120-Ile201B-Ile424-Ala433 (799) AAGCTGCAGGCCCAGTTCGGCAACAAGACC Consensus (841) AAGCTGCAGGCCCAGTTCGGCAACAAGACC Leu122-Ser199-Tryp427-Gly431 (841) ATCGTGTTCAAGCAGAGCAGCGGCGGCGAC Val127-Asn195-Arg426-Gly431 (871) ATCGTGTTCAAGCAGAGCAGCGGCGGCGAC Val120-Thr202-Ile424-Ala433 (829) ATCGTGTTCAAGCAGAGCAGCGGCGGCGAC Leu122-Ser199-Arg426-Lys432 (841) ATCGTGTTCAAGCAGAGCAGCGGCGGCGAC Leu122-Ser199-Arg426-Gly431 (841) ATCGTGTTCAAGCAGAGCAGCGGCGGCGAC Lys121-Val200-Asn425-Lys432 (835) ATCGTGTTCAAGCAGAGCAGCGGCGGCGAC Val120-Ile201-Ile424-Ala433 (829) ATCGTGTTCAAGCAGAGCAGCGGCGGCGAC Val120-Ile201B-Ile424-Ala433 (829) ATCGTGTTCAAGCAGAGCAGCGGCGGCGAC Consensus (871) ATCGTGTTCAAGCAGAGCAGCGGCGGCGAC Leu122-Ser199-Tryp427-Gly431 (871) CCCGAGATCGTGATGCACAGCTTCAACTGC Val127-Asn195-Arg426-Gly431 (901) CCCGAGATCGTGATGCACAGCTTCAACTGC Val120-Thr202-Ile424-Ala433 (859) CCCGAGATCGTGATGCACAGCTTCAACTGC Leu122-Ser199-Arg426-Lys432 (871) CCCGAGATCGTGATGCACAGCTTCAACTGC Leu122-Ser199-Arg426-Gly431 (871) CCCGAGATCGTGATGCACAGCTTCAACTGC Lys121-Val200-Asn425-Lys432 (865) CCGGAGATCGTGATGCACAGCTTCAACTGC Val120-Ile201-Ile424-Ala433 (859) CCCGAGATCGTGATGCACAGCTTGAACTGC Val120-Ile201B-Ile424-Ala433 (859) CCCGAGATCGTGATGCACAGCTTCAACTGC Consensus (901) CCCGAGATCGTGATGCACAGCTTCAACTGC Leu122-Ser199-Tryp427-Gly431 (901) GGCGGCGAGTTCTTCTACTGCAACAGCACC Val127-Asn195-Arg426-Gly431 (931) GGCGGCGAGTTCTTCTACTGCAACAGCACC Val120-Thr202-Ile424-Ala433 (889) GGCGGCGAGTTCTTCTACTGCAACAGCACC Leu122-Ser199-Arg426-Lys432 (901) GGCGGCGAGTICTTCTACTGCAACAGCACC (901) GGCGGGAGTTCTTCTACTGCAACAGCACC Leu122-Ser199-Arg426-Gly431 Lys121-Val200-Asn425-Lys432 (895) GGCGGCGAGTTCTTCTACTGCAACAGCACC Val120-Ile201-Ile424-Ala433 (889) GGCGGGGAGTTCTTCTACTGCAACAGCACC Val120-Ile201B-Ile424-Ala433 (889)GGCGGCGAGINICITICTACTGCAACAGCACC Consensus (931)GGCGGCGAGTTCTTCTACTGCAACAGCACC Leu122-Ser199-Tryp427-Gly431 (931) CAGCTGTTCAAGAGCACGTGGAAGAGACC Val127-Asn195-Arg426-Gly431 (961) CAGCTGTTCAACAGCACCTGGAACACACC (919) CAGCTGTTCAACAGCACCTGGAACACACC Val120-Thr202-Ile424-Ala433 Leu122-Ser199-Arg426-Lys432 (931)CAGCTGTTCAACAGCACCTGGAACACACC Leu122-Ser199-Arg426-Gly431 (931)CAGCTGTTCAACAGCACCTGGAACAACACC Lys121-Val200-Asn425-Lys432 (925)CAGCIGITCAACAGCACGTGGAACAACACC Val120-Ile201-Ile424-Ala433 CAGCTGTTCAACAGCACCTGGAACACACC (919) Val120-Ile201B-Ile424-Ala433 CAGCTGATCAACAGCACCTGGAACAACACC (919) Consensus (961) CAGCTGTTCAACAGCACCTGGAACAACACC 991 1020 Leu122-Ser199-Tryp427-Gly431 (961) ATEGGECCAACAACACCAACGGCACCATC Val127-Asn195-Arg426-Gly431 (991) ATCGGCCCAACACACCAACGGCACCATC Val120-Thr202-Ile424-Ala433 (949) ATCGGCCCAACACACCAACGGCACCATC Leu122-Ser199-Arg426-Lys432 (961) ATCGCCCCAACACACCAACGCCACCATC
(961) ATCGCCCCCAACACACCACCAACGCCACCATC Leu122-Ser199-Arg426-Gly431 Lys121-Val200-Asn425-Lys432 (955) ATCGGCCCAACACACCAACGGCACCATC Val120-Ile201-Ile424-Ala433 (949) ATCGGCCCAACACACCAACGGCACCATC Val120-Ile201B-Ile424-Ala433 (949) ATCGGCCCAACAACACCAACGGCACCATC Consensus (991) ATCGGCCCCAACAACACCAACGGCACCATC Leu122-Ser199-Tryp427-Gly431 (991) ACCCTGCCCTGCCGCATCAAGCAGATCATC

```
(1021) ACCCTGCCCTGCCGCATCAAGCAGATCATC
 Val127-Asn195-Arg426-Gly431
                                (979) ACCCTGCCTGCCGCATCAAGCAGATCATC
 Val120-Thr202-Ile424-Ala433
                                (991) ACCCTGCCTGCCGCATCAAGCAGATCATC
 Leu122-Ser199-Arg426-Lys432
                                (991) ACCCTGCCTGCCGCATCAAGCAGATCATC
 Leu122-Ser199-Arg426-Gly431
                                (985) ACCCTGCCCTGCCGCATCAAGCAGATCATC
 Lys121-Val200-Asn425-Lys432
 Val120-Ile201-Ile424-Ala433
                                (979) ACCCTGCCCTGCCGCATCAAGCAGATCATC
                                (979) ACCCTGCCCTGCCGCATCAAGCAGATCATC
Val120-Ile201B-Ile424-Ala433
                               (1021) ACCCTGCCCTGCCGCATCAAGCAGATCATC
                   Consensus
Leu122-Ser199 Tryp427-Gly431
                               (1021) AACCGCTGGGGCGGCAAGGCCATGTACGCC
                               (1051) AACCGCGGCGGCGGCAAGGCCATGTACGCC
 Val127-Asn195-Arg426-Gly431
                               (1009) -----GCCGCC--GCCATGTACGCC
 Val120-Thr202-Ile424-Ala433
                               (1021) AACCGCGGCGGCAACAAGGCCATGTACGCC
 Leu122-Ser199-Arg426-Lys432
 Leu122-Ser199-Arg426-Gly431
                               (1021) AACCGCGGCAGCGCAAGGCCATGTACGCC
                               (1015) AAC-----GCCCCCAAGGCCATGTACGCC
 Lys121-Val200-Asn425-Lys432
 Val120-Ile201-Ile424-Ala433
                                      -----GGCGGC---GCCATGTACGCC
                                     -----GGCGGC---GCCATGTACGCC
Val120-Ile201B-Ile424-Ala433
                               (1009)
                               (1051) AACCGC G GGCGGCAAGGCCATGTACGCC
                   Consensus
                                      1081
                                                                1110
                               (1051) CCCCCCATCCGCGCCAGATCCGCTGCAGC
Leu122-Ser199 Tryp427-Gly431
                                      CCCCCATCCGCGGCCAGATECGCTGCAGC
 Val127-Asn195-Arg426-Gly431
                               (1081)
                                      CCCCCATCCGCGGCCAGATCCGCTGCAGC
                               (1027)
 Val120-Thr202-Ile424-Ala433
 Leu122-Ser199-Arg426-Lys432
                               (1051)
                                      CCCCCATCCGCGCCAGATCCGCTGCAGC
 Leu122-Ser199-Arg426-Gly431
                               (1051)
                                      CCCCCATCCGCGGCCAGATCCGCTGCAGC
                                      CCCCCATCCGCGGCCAGATCCGCTGCAGC
                               (1039)
 Lys121-Val200-Asn425-Lys432
                               (1027)
                                      CCCCCCATCCGCGGCCAGATCCGCTGCAGC
 Val120-Ile201-Ile424-Ala433
                               (1027) CCCCCCATCCGCGGCCAGATCCGCTGCAGC
Val120-Ile201B-Ile424-Ala433
                   Consensus
                               (1081) CCCCCCATCCGCGGCCAGATCCGCTGCAGC
                               (1081) AGCAACATCACCGGCCTGCTGCTGACCCGC
Leu122-Ser199 Tryp427-Gly431
 Val127-Asn195-Arg426-Gly431
                               (1111)
                                      AGCAACATGACCGGCCTGCTGCTGACCCGC
 Val120-Thr202-Ile424-Ala433
                               (1057)
                                      AGEANGANDACEGGOOTEGETEGTEACCOGC
                                      AGCAACA I GAUGGGCCII GUITGUTGACCEGC
 Leu122-Ser199-Arg426-Lys432
                               (1081)
                                      AGCAACATCACCGGCCTGCTGCTGACCCGC
                               (1081)
 Leu122-Ser199-Arg426-Gly431
Lys121-Val200-Asn425-Lys432
                               (1069)
                                      AGCAACATCACCGGCCTGCTGCTGACCCGC
                               (1057)
                                      AGe: V1/w; v1/e/GGGGGGH; Geric erite; CerceC
 Val120-Ile201-Ile424-Ala433
                               (1057)
                                     AGGAAGATCAGCGGCCTGGRGGRGAGGCGC
Val120-Ile201B-Ile424-Ala433
                   Consensus
                               (11111)
                                     AGCAACATCACCGGCCTGCTGCTGACCCGC
                                      GAEGGEGEAAGGAGATCAGGAAGACCACC
Leu122-Ser199 Tryp427-Gly431
                               (11111)
                                      GACGGGGGAAGGAGATCAGCARCACCACC
 Val127-Asn195-Arg426-Gly431
                               (1141)
                               (1087)
                                      GACCCCCCAACGAGATCAGCAACACCACC
 Val120-Thr202-Ile424-Ala433
                                      GACGGCCCAAGGAGATCAGCAACACCAGC
 Leu122-Ser199-Arg426-Lys432
                               (11111)
                                     CASCOCCONTROCACATROACOARSOACOA
 Leu122-Ser199-Arg426-Gly431
                               (11111)
                               (1099)
                                      GACGGCGGCAAGGAGATCAGCAACACCACC
 Lys121-Val200-Asn425-Lys432
                               (1087)
                                      GACGGGGCAAGGAGATCAGCAACACCACC
 Val120-Ile201-Ile424-Ala433
Val120-Ile201B-Ile424-Ala433
                               (1087)
                                      GACGECGECAAGGAGATCAGEAACACCACC
                               (1141)
                                      GACGGCGCAAGGAGATCAGCAACACCACC
                   Consensus
                                      1171
                                      GAGATETTCCGCCCGGCGGCGGCGACATG
Leu122-Ser199 Tryp427-Gly431
                               (1141)
                                      GAGATOTTCCGCCCGGCGGCGACATG
 Val127-Asn195-Arg426-Gly431
                               (1171)
                               (1117)
                                      GAGATETTCCGCCCCGGCGGCGGCGACATG
 Val120-Thr202-Ile424-Ala433
                               (1141)
                                      GAGATETTECGCCCCGGCGCGCGCGACATG
Leu122-Ser199-Arg426-Lys432
Leu122-Ser199-Arg426-Glv431
                               (1141)
                                      GAGATETTEEGCCCCGGCGGCGGCGACATG
Lys121-Val200-Asn425-Lys432
                               (1129)
                                      GAGATCTTCCGCCCCGGCGCGCGCGACATG
                                     GAGATETTECGCCCGGCGGCGGCGACATG
 Val120-Ile201-Ile424-Ala433
                               (1117)
Val120-Ile201B-Ile424-Ala433
                               (1117) GAGATETTCCGCCCCGGCGCGCGCGACATG
```

1201 1201 1202 1203 1203 1204 1204 1204 1207	Consensus	(1171)	GAGATCTTCCGCCCCGGCGCGCGACATG
Leu122-Ser199 Tryp427-G1y431 (1171) GCGTATACTSC CACCACCT COMB Val120-Thr202-Ile424-Ala433 (1171) GCGTATACTSC CACCACCT ACAD Val120-Thr202-Ile424-Ala433 (1171) GCGTATACTSC CACCACCTS ACAD Val120-Ile201-Ile424-Ala433 (1171) GCGTATACTSC CACCACCTS ACAD Val120-Thr202-Ile424-Ala433 (1171) GCGTATACTSC CACCACCTS ACAD Val120-Thr202-Ile424-Ala433 (1171) TATACTSC CACCACCTS CACCACCTS ACAD Val120-Thr202-Ile424-Ala433 (1171) TATACTSC CACCACCTS CACCACCTS CACCACCTS Val120-Ile201-Ile424-Ala433 (1171) TATACTSC CACCACCTS CACCACCTS Val120-Ile201-Ile424-Ala433 (1171) TATACTSC CACCACCTS CACCACCACCTS CACCACCTS CACCACCACCTS CACCACCTS CACCACCTS CACCACCTS CACCACCTS CACCACCTS CACCACCACCTS CACCACCTS CACCACCTS CACCACCTS CACCACCACCACCCTS CACCACCACCACCACCACCCTS CACCACCACCACCCTS CACCACCACCACCACCACCACCCTS CACCACCACCACCACCACCACCACCACCACCACCACCAC	00561545	(
Val127-Asn195-Arq426-Giy431	Leu122-Ser199 Trvp427-Glv431	(1171)	
Val120-Thr202-Ile424-Ala433			
Leul22-Ser199-Arg426-Gly431 (1171) COURT			
Leul22-Ser199-Arg426-Ciy431 (1171) STARAS SECTION TABLE Val120-Ile201-Ile424-Ala433 (1147) GGGGCAACTGGCGAGCGTTTALAG CGGGCAACTGGCGAGCGTTTALAG CGGGCAACTGGCGAGCGCACTGTACAG 1231 1260 Leul22-Ser199 Tryp427-Gly431 (1201) TACAAGGTGGTGAACTGGCGCAGCGCACTGTACAG CGGCCAACTGGCGAGCACTGTACAG CGGCCAACTGGCGCGCACCTGTACAG CGGCCAACTGGCGCAGCCATTACAGCCTGGCC Val127-Asn195-Arg426-Gly431 (1201) TACAAGGTGGTGAACTGCCCCGCGCACCTGGC Val120-Thr202-Ile424-Ala433 (1177) TACAAGGTGGTGAACATCGGCCCAGCCCTGGC Lys121-Val200-Asn425-Lys432 (189) TACAAGGTGGTGAACATCGGCCCTGGC Val120-Ile201-Ile424-Ala433 (1177) TACAAGGTGGTGAACATCGACCCCTGGC Val120-Ile201-Ile424-Ala433 (1177) TACAAGGTGGTGAACATCGACCCCTGGC Val120-Ile201-Ile424-Ala433 (1177) TACAAGGTGGTGAACATCGACCCCTGGC Val120-Thr202-Ile424-Ala433 (1177) TACAAGGTGGGAACATCGACCCCTGGC Val120-Thr202-Ile424-Ala433 (1201) TACAAGGTGGTGAACATCGACCCCTGGGC Val120-Thr202-Ile424-Ala433 (1201) TACAAGGTGGTGAACATCGACCCCTGGGC Val120-Thr202-Ile424-Ala433 (1201) TACAAGGTGGTGAACATCGACCCCTGGGC Val120-Thr202-Ile424-Ala433 (1201) TACAAGGTGGGAACACGCCGGTG Val120-Thr202-Ile424-Ala433 (1201) TACAAGGTGGGAACACGCCGGTG Val120-Thr202-Ile424-Ala433 (1201) TACAAGGTGGGAACACGCCGGGTG Val120-Thr202-Ile424-Ala433 (1201) TACAAGGTGGGAACGCCGGGTG Val120-Thr202-Tle424-Ala433 (1201) TACAAGGTGGGAACGCCGGGTG Val120-Thr202-Tle424-Ala433 (1201) TACAAGGTGGGAACGCCGGGTG Val120-Thr202-Tle424-Ala433 (1201) TACAAGGTGGGAACGCCGGGTG Val120-Thr202-Tle424-Ala433 (1201) TACAAGGTGGGAACGCCGCGGTG (1201) TACAAGGTGGGAACGCCGGGTG (1201) TACAAGGTGGGAACGCCGCGGTG (1201) TACAAGGTGGGAACGCGCGGTG (1201) TACAAGGTGGGAACGCGCGGTG (1201) TACAAGGTGGGAACGCGCGGTG (1201) TACAAGGTGGGAACGCGCGGTG (1201) TACAAGGTGGGAACGCGCGGTGGGGGGAACAGGCCGGGTGGAACGCGCGGGGAACAGGCCGAACGCGCGGGGAACAGGCCGAACGACG			
Lys121-Val200-Asn425-Lys432			
Val120-Ile201-Ile424-Ala433 (1147) CGCGACAACTGGCCAGCCAGTTACAS CONSENSUS (1201) CGCGACAACTGGCCGCAGCCAGCTTACAS 1231 1260 Leu122-Ser199 Tryp427-Gly431 (1201) TACAGGTGCGAGCAGCTGACGCAGCTGACAG Val127-Asn195-Arg426-Gly431 (1231) TACAGGTGCGAGCAGCTGGCGCAGCCTGGGC Val127-Asn195-Arg426-Gly431 (1201) TACAGGTGCGAGCAGCTGGGCCAGCCTGGGC Leu122-Ser199-Arg426-Gly431 (1201) TACAGGTGCTGAGAGTCAGCCCTGGGC Leu122-Ser199-Arg426-Gly431 (1201) TACAGGTGCTGAGAGTCAGCCCTGGGC Leu122-Ser199-Arg426-Gly431 (1201) TACAGGTGCTGAGAGTCAGCCCTGGGC Val120-Ile201-Ile424-Ala433 (1177) TACAGGTGCTGAGAGTCAGCCCTGGGC Val120-Ile201-Ile424-Ala433 (1177) TACAGGTGCTGAGAGTCGAGCCCCTGGGC Leu122-Ser199 Tryp427-Gly431 (1201) TACAGGTGCTGAGAGTCGAGCCCCTGGGC Leu122-Ser199 Arg426-Gly431 (1201) TACAGGTGCTGAGAGTCGAGCCCCTGGGC Leu122-Ser199-Arg426-Gly431 (1201) TACAGGTGCTGAGAGTCGAGCCCCTGGGC Leu122-Ser199-Arg426-Gly431 (1201) GTGGCCCCACCAAGGCCAAGGCCGGTG Leu122-Ser199-Arg426-Gly431 (1201) GTGGCCCCACCAAGGCCAAGGCCGGTG Leu122-Ser199-Arg426-Gly431 (1201) GTGGCCCCACCAAGGCCAAGGCCGGTG Leu122-Ser199-Arg426-Gly431 (1201) GTGGCCCCACCAAGGCCAAGGCCGGGTG Leu122-Ser199-Arg426-Gly431 (1201) GTGGCCCCACCAAGGCCAAGGCCGGGTG Leu122-Ser199-Arg426-Gly431 (1201) GTGGCCCCACCAAGGCCAAGGCCGGGTG Leu122-Ser199-Arg426-Gly431 (1201) GTGGCCCCACCAAGGCCAAGGCCGGGTG Leu122-Ser199-Arg426-Gly431 (1201) GTGGCCCCACCAAGGCCAAGGCCGGCTG Leu122-Ser199-Arg426-Gly431 (1201) GTGGCCCCACCAAGGCCAAGGCCGGCTG Leu122-Ser199-Arg426-Gly431 (1201) GTGGCCCCACCAAGGCCGAGCGCGCTGCCTG Leu122-Ser199-Arg426-Gly431 (1201) GTGGCCCCACCAAGGCCGAGCGCGCTGCCTG Leu122-Ser199-Arg426-Gly431 (1201) GTGGCCCCACCAAGGCCGAGCGCGCTGCCTG Leu122-Ser199-Arg426-Gly431 (1201) GTGGCCCCACCAGGCGCGCTGCCTGCCTG Leu122-Ser199-Arg426-Gly431 (1201) GTGGCCCCACCAGGCGCGCTGCCCTG Leu122-Ser199-Arg426-Gly431 (1201) GTGGCCCCACCAGGCGCGCTGCCCTG Leu122-Ser199-Arg426-Gly431 (1201) GTGGCCCCACCAGCGCGCGTGCCCTG Leu122-Ser199-Arg426-Gly431 (1201) GTGGCCCCACAGGCCGCTGCCCTGCCCC			
Val120-Ile201B-Ile424-Ala433			
Consensus			
Leul22-Ser199 Tryp427-Gly431	Consensus		
Val127-Asn195-Arg426-Gly431			1231 1260
Val127-Asn195-Arg426-Gly431	Leu122-Ser199 Tryp427-Gly431	(1201)	TACAAGGTGGTGAAGATCGAGCCCCTGGGC
Val120-Thr202-Tle424-Ala433 (1177)	Val127-Asn195-Arg426-Gly431	(1231)	TACAAGGTGGTGAAGATCGAGCCCCTGGGC
Leu122-Ser199-Arg426-Lys432		(1177)	
Leu122-Ser199-Arg426-Gly431	Leu122-Ser199-Arg426-Lys432	(1201)	
Lys121-Val200-Asn425-Lys432 (1189) RCARGOTS-TARGATCGRGCCCTGGGG Val120-Ile201B-Ile424-Ala433 (1177) TACARGOTS-TARGATCGRGCCCTGGGG (1291) TACARGOTS-TARGATCGRGCCCTGGGG (1291) TACARGOTS-TARGATCGRGCCCCTGGGG (1291) TACARGOTS-TARGATCGRGCCCCTGGGG (1291) TACARGOTS-TARGATCGRGCCCCTGGGG (1291) TACARGOTS-GARGATCGRGCCCCTGGGG (1291) TACARGOTS-GARGATCGRGCCCCTGGGC (1291) TACARGOTS-GARGATCGRGCCCCTGGGC (1291) TACARGOTS-GARGATCGRGCCCCTGGGC (1291) TACARGOTS-GARGATCGRGCCCCTGGGC (1291) TACARGOTS-GARGATCGRGCCCCCTGGGC (1291) TACARGOTS-GARGATCGRGCCCCCTGGGC (1291) TACARGOTS-GARGATCGRGCCCAGCGGTG (1291) TACARGOTS-GARGATCGRGCCAGCGCGTG (1291) TACARGOTS-GARGATCGRGCCAGCGCGTG (1291) TACARGOTS-GARGATCGRGCCAGCGCGTG (1291) TACARGOTS-GARGATCGRGCCCAGCGGTG (1291) TACARGOTS-GARGATCGRGCCAGCGCGTG (1291) TACARGOTS-GARGATCGRGCCAGCGCGTG (1291) TACARGOTS-GARGATCGRGCCAGCGCGTG (1291) TACARGOTS-GARGATCGRGCCAGCGCGTG (1291) TACARGOTS-GARGATCGRGCCAGCGCGTG (1291) TACARGOTS-GARGATCGRGCCAGCGCGTG (1291) TACARGOTS-GARGATCGRCCAGCGCGTG (1291) TACARGOTS-ARGGCCAGCGCGTG (1291) TACARGOTS-GARGATCGRCCAGCGCGTG (1291) TACARGOTS-GARGATCGRCCAGCGCGTG (1291) TACARGOTS-GARGATCGGCCAGCGGTG (1291) TACARGOTS-GARGATCGGCCAGCGTG (1291) TACARGOTS-GARGATCGGCCAGCGGTG (1291) TACARGOTS-GARGATCGGCCAGGGCGTG (1291) TACARGOTS-GARGATCGGCCAGGGCGTG (1291) TACARGOTS-GARGATCGGCCAGGGCCGGGGGGTG (1291) TACARGOTS-GARGATCGGCCAGGGCGGGGGGGGGGGGGG	Leu122-Ser199-Arg426-Gly431	(1201)	
Val120-Ile201-Ile424-Ala433		(1189)	7-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1
Val120-Ile201B-Ile424-Ala433 Consensus			
Consensus	Val120-Ile201B-Ile424-Ala433	(1177)	TACAAGGTGGTGAACATCGAGCCCCTGGGC
1261 1290 1291	Consensus	(1231)	TACAAGGTGGTGAAGATCGAGCCCCTGGGC
Val127-Asn195-Arg426-Gly431			
Val127-Asn195-Arg426-Gly431	Leu122-Ser199 Tryp427-Gly431	(1231)	GTGGGGGCCACCAAGGGGAAGGGGGGGGTG
Leu122-Ser199-Arg426-Lys432 (1231) GTGGCCC ACAAGGCCAAGGCGGGGGGGGGGACLYS121-Val200-Ass425-Lys432 (1219) GTGGCCCCAACAAGGCCAAGAGCCAAGCCAAGGCCAAGACCAAGCCAAGGCCAAGACCAAGCACAAGCCAAGCAAGCCAAGCCAAGCAAGCCAAGC	Val127-Asn195-Arg426-Gly431	(1261)	
Leu122-Ser199-Arg426-Lys432 (1231) GTGGCCCCACCAAGGCCAAGCCCGGCTG Lys121-Val200-Asn425-Lys432 (1219) GTGGCCCCAAGGCCAAGGCCAAGGCCAGGCTG Val120-Tle201-Tle424-Ala433 (1207) GTGGCCCCAACGCCAAGGCCAAGGCCGCTG Val120-Tle201B-Ile424-Ala433 (1207) GTGGCCCCACCAAGGCCAAGGCCGCGTG Consensus (1261) GTGGCCCCACCAAGGCCAAGGCCGCGTG Val127-Asn195-Arg426-Gly431 (1261) GTGGCCCCACCAAGGCCAAGGCCGCGTG Val120-Thr202-Tle424-Ala433 (1297) GTGGCCCCACCAAGGCCAAGGCCGCGTG Val120-Thr202-Tle424-Ala433 (1291) GTGGGCCCCACCAAGGCCAAGGCCCGCGTGACCTG Val120-Thr202-Tle424-Ala433 (1291) GTGGGCCCCACCAAGGCCAAGGCCGCGTGACCTG Leu122-Ser199-Arg426-Gly431 (1261) GTGCAGCGCAAAGCGGCGCGTGACCTG Val120-Tle201-Tle424-Ala433 (1237) GTGCAGCGAAAAGCGGCGCGTGACCTG Val120-Tle201B-Ile424-Ala433 (1237) GTGCAGCGAAAAGCGGCGCGTGACCCTG Val120-Thr202-Tle424-Ala433 (1237) GTGCAGCGAAAAGCGGCGCGTGACCCTG Val120-Thr202-Tle424-Ala433 (1237) GTGCAGCGAAAAGCGGCGCGTGACCCTG Val120-Thr202-Tle424-Ala433 (1237) GTGCAGCGAAAAGCGGGCCGTGACCCTG CGGCCAAGAACGGGCGCGTGACCCTG Val120-Thr202-Tle424-Ala433 (1237) GTGCAGCGAAAAGCGGGCCCGTGACCCTG Val120-Thr202-Tle424-Ala433 (1237) GTGCAGCGAAAAGCGGGCCCGTGACCCTG CGCCAAGAACACGGCGCCTGACCCTG CGCCAACAAACACGGCCCCTGACCCTG CGCCAACAAACACGGCCCCTGACCCTG CCAACAAACACGCGCCCTGACCCTG CCAACAACACACGGCCCCTGACCCTG CCAACAACACACGCCCCCTGACCCTG CCAACAACACACGCCCCCTGACCCTG CCAACAACACACGCCCCCTGACCCTG CCAACAACACACGCCCCCTGACCCTG CCAACAACACACACACACACACACACACACACACACAC	Val120-Thr202-Ile424-Ala433	(1207)	GTGGGCCCGACCAAGGCCGAAGGGCCGCGTG
Leu122-Ser199-Arg426-Gly431 (1231) GTGGCCCCACCAGGCCAGGCCGGGTG Lys121-Val200-Assn425-Lys432 (1219) GTGGCCCCACCAAGGCCAAGGCCAGGCGGGTG Val120-Ile201B-Ile424-Ala433 (1207) GTGGCCCCACCANGGCCAAGGCCGGGTG CACCANGGCCAAGGCCCAAGGCCGGGTG CACCANGGCCAAGGCCCAAGACCGCCCCAAGACCACAAGGCCCAAGACCACC	Leu122-Ser199-Arg426-Lys432	(1231)	
Lys121-Val200-Asn425-Lys432	Leu122-Ser199-Arg426-Gly431	(1231)	
Val120-Ile201-Ile424-Ala433	Lys121-Val200-Asn425-Lys432	(1219)	GTGGCCCCCACCAAGGCCAAGGGCCGCGTG
Consensus	Val120-Ile201-Ile424-Ala433	(1207)	GTGGCCCCCACCAAGGCCAAGGGCCGCGTG
Consensus	Val120-Ile201B-Ile424-Ala433	(1207)	
Leu122-Ser199 Tryp427-Gly431	Consensus	(1261)	
Val127-Asn195-Arg426-Gly431			1291 1320
Val120-Thr202-Ile424-Ala433	Leul22-Serl99 Tryp427-Gly431	(1261)	GTGCAGCGCGAGAAGCGCGCGCGTGACCCTG
Leu122-Ser199-Arg426-Lys432 (1261) GTGCAGCGCACATGCGCGCCATGACCATG Lys121-Val200-Asn425-Lys432 (1249) GTGCAGCGCCATGAAGCGCGCCATGACCATG Val120-Ile201-Ile424-Ala433 (1237) GTGCACCGCATGAAGCGCGCCATGACCATG Val120-Ile201B-Ile424-Ala433 (1237) GTGCACCGCATGAAGCGCGCCATGACCATG Consensus (1291) GTGCAGCGCAGAAGCGCGCCGTGACCATG Val127-Asn195-Arg426-Gly431 (1291) GGCGCATGACCATGCCCCCCCCCCCCCCCCCCCCCCCCC	Val127-Asn195-Arg426-Gly431	(1291)	GTGCAGCGCGAGAAGCGCGCGTGACCCTG
Leu122-Ser199-Arg426-Gly431 (1261) GTGCAGCG AGAAGGGGGCGTGACCTG Lys121-Val200-Asn425-Lys432 (1249) GTGCAGCGCAGAAGGGGCGCTGACCTG Val120-Ile201-Ile424-Ala433 (1237) GTGCAGCGCAGAAGGGGCGCGTGACCTG Consensus (1291) GTGCAGCGCAGAAGGGGCGCGTGACCCTG 1321 - 1350 Leu122-Ser199 Tryp427-Gly431 (1291) GGGGCAGAAGGGGGCGCGTGACCCTG 1321 - 1350 Leu122-Ser199-Arg426-Gly431 (1321) GGGGCATGTCCTGGGGGCC Val120-Thr202-Ile424-Ala433 (1267) GGGGCATGTCTGGGGGCC Leu122-Ser199-Arg426-Gly431 (1291) GGGGCATGTCTGGGGGGC Val120-Ile201-Ile424-Ala433 (1267) GGGGCATGTTCTGGGGGGCC Val120-Ile201-Ile424-Ala433 (1267) GGGGCATGTTCTGGGGGGCC Val120-Ile201-Ile424-Ala433 (1267) GGGCCATGTTCTGGGCGCC Consensus (1321) GGCGCATGTTCTGGGCTTCTGGGCGCC Val120-Ile201-Ile424-Ala433 (1267) GGCGCATGTTCTGGGCTTCTGGGCGCC Consensus (1321) GGCGCATGTTCTTGGGCTTCTTGGGCGCC Val120-Thr202-Ile424-Ala433 (1267) GGCGCATGTTCTTGGGCTTCTTGGGCGCC Consensus (1321) GCCGCATGTTCTTGGGCTTCTTGGGCGCC Val120-Thr202-Ile424-Ala433 (1267) GCCGCATGTTCTTGGGCTTCTTGGGCGCCC Val120-Thr202-Tle424-Ala433 (1267) GCCGCATGTTCTTGGGCTTCTTGGGCGCCCCCCCCCCCC	Val120-Thr202-Ile424-Ala433	(1237)	GTGGAGCGCGAGAAGCGGGGGGGGTGACCCTG
Lys121-Val200-Asn425-Lys432 (1249) GTGCAGCG GAGAAGCGCCCGTGACCCTG Val120-Ile201-Ile424-Ala433 (1237) GTGCAGCG GAGAAGCGCCCCTGACCCTG Val120-Ile201B-Ile424-Ala433 (1237) GTGCAGCGGAGAAGCGCCCCTGACCCTG Consensus (1291) GTGCAGCGGAGAAGCGCCCCTGACCCTG 1321 - 1350 Leul22-Ser199 Tryp427-Gly431 (1291) GGCGCATGTCTGACCCTG Val127-Asn195-Arg426-Gly431 (1321) GGCGCATGTTCTGGGCTCC Val120-Thr202-Ile424-Ala433 (1267) GGCGCATGTTCTGGGCTCCTGGGCCC Leul22-Ser199-Arg426-Gly431 (1291) GGCGCATGTTCTGGGCTCCTGGGCCCC Val120-Ile201-Ile424-Ala433 (1267) GGCGCATGTTCTTGGGCTCCTGGGCCCC Val120-Ile201B-Ile424-Ala433 (1267) GGCGCATGTTCTTGGGCTTCCTGGGCCCC Val120-Ile201B-Ile424-Ala433 (1267) GGCGCATGTTCTTGGGCTTCCTGGGCCCC Val127-Asn195-Arg426-Gly431 (1321) GCCGCATGTTCCTGGGCTTCCTGGGCCCC Val127-Asn195-Arg426-Gly431 (1351) GCCGCATGTTCCTGGGCTCCCCCCCCCCCCCCCCCCCCC	Leu122-Ser199-Arg426-Lys432	(1261)	GTGCAGCGCGAGAAGCGCGCGCGTGACCCTG
Val120-Ile201-Ile424-Ala433 (1237) GTGCAGGGGAGAGGGGGGGGGGGGGGGGGGGGGGGGGGG	Leu122-Ser199-Arg426-Gly431	(1261)	GTGCAGCGCCAGAAGCGCGCGCGTGACCCTG
Val120-Ile201-Ile424-Ala433 (1237) GTGCAGGGGAGAGGGGGGGGGGGGGGGGGGGGGGGGGGG	Lys121-Val200-Asn425-Lys432	(1249)	GTGCAGCGCGAGAAGCGCGCCGTGACCCTG
Vall20-Ile201B-Ile424-Ala433	Val120-Ile201-Ile424-Ala433	(1237)	
1321	Vall20-Ile201B-Ile424-Ala433	(1237)	
Leu122-Ser199 Tryp427-Gly431 (1291) GGCGCATGTTCCTGGGCGCCCCCCCCCCCCCCCCCCCCC	Consensus	(1291)	GTGCAGCGCGAGAAGCGCGCGTGACCCTG
Val127-Asn195-Arg426-Gly431 (1321) GGCECATAR STEGGCTIC GGCGCCCCCCCCCCCCCCCCCCCCCCCCCCCCC			1321 - 1350
Val120-Thr202-Ile424-Ala433 (1267) GGGECATGIT CIGGGGTTG TEGGGGCC Leu122-Ser199-Arg426-Lys432 (1291) GGGCATGIT CIGGGGTTG TEGGGGCC Leu122-Ser199-Arg426-Gly431 (1291) GGGCATGIT CIGGGGTTC TEGGGGCC Lys121-Val200-Asn425-Lys432 (1279) GGGCATGIT CIGGGGTTC TEGGGCCC Val120-Ile201-Ile424-Ala433 (1267) GGGGCATGIT CIGGGCTTC TEGGGCCC Val120-Ile201B-Ile424-Ala433 (1267) GGGGCCATGIT CIGGGCTTC TEGGGCCC (1321) GGCGCCATGIT CIGGGCTTC TEGGGCGCC 1351 1380 Leu122-Ser199 Tryp427-Gly431 (1321) GCCGGCAGCACATGGGCGCCACCACCACCACCACCACCACCACCACCACCACC		(1291)	GGCGCCATGTTCCTGGGCCCC
Leu122-Ser199-Arg426-Lys432 (1291) GGGCTGCTTCTGGGCGCC Lys121-Val200-Asn425-Lys432 (1291) GGGCTAGTTCTGGGCTTCTTGGGCGCC Val120-Ile201-Ile424-Ala433 (1267) GGGCTAGTTCTGGGCTTCTGGGCGCC (1321) GGCGCATGTTCTGGGCTTCCTGGGCGCC (1321) GGCGCATGTTCTGGGCTTCCTGGGCGCC (1321) GGCGCATGTTCTGGGCTTCCTGGGCGCC (1321) GGCGCCATGTTCCTGGGCTGCC (1321) GGCGCCATGTTCCTGGGCTGCC (1321) GGCGCCATGTTCCTGGGCTGCC (1321) GGCGCCATGTTCCTGGGCTGCC (1321) GGCGCCATGTTCCTGGGCGCC (1321) GGCGCCATGTTCCTGGGCGCC (1321) GGCGCCATGTTCCTGGGCGCC (1321) GGCGCCATGTTCCTGGGCGCC (1321) GGCGCCATGTTCCTGGGCGCC (1321) GCCGCAGCACATGGCGCCCACCACGTG (1321) GCCGCAGCACATGGCGCCCCACCACGTG (1321) GCCGCAGCACATGGCCCCCCACCACGTG (1321) GCCGCAAGCACATGGCCCCCCCACCACGTG (1321) GCCGCAAGCACATGGCCCCCCCACCACGTG (1321) GCCGCAAGCACATGGCCCCCCCACCACGTG (1321) GCCGCAAGCACATGGCCCCCCCCACCACGTG (1321) GCCGCAAGCACATGGCCCCCCCCACCACGTG (1321) GCCGCAAGCACATGGCCCCCCCCCACCACGTG (1321) GCCGCCACCACCACCACGTG (1321) GCCGCCACCACCACCACCACCACCACCACCACCACCACCA	Val127-Asn195-Arg426-Gly431	(1321)	GGCGCCATGTECCTGGGCTTCCTGCGGGCC
Leu122-Ser199-Arg426-Gly431 (1291) GGCGCATGTT TGGGGTTCCTGGGGGCC Lys121-Val200-Asn425-Lys432 (1279) GGCGCATGTT TGGGGTTCCTGGGGGCC Val120-Ile201-Ile424-Ala433 (1267) GGCGCATGTT TGGGGTTCCTGGGGGCC (1321) GGCGCATGTT TGGGGTTCCTGGGGGCC (1321) GGCGCCATGTTCCTGGGCGCC (1321) GCCGCCAGCACATGTCGCCCCCAGCACATGTCCTGGCGCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCCAGCACATGTCCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCCAGCACATGTCCCCCCCAGCACATGTCCCCCCCAGCACATGTCCCCCCCAGCACATGTCCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCCAGCACATGTCCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCCAGCACATGTCCCCCCCAGCACATGTCCCCCCCAGCACATGTCCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCAGCACAACATGCCCCCCAGCACAACAACAACAACAACAACAACAACAACA		(1267)	GGCGCCATGCCCCGGCCCCCCCCCCCCCCCCCCCCCCCC
Leu122-Ser199-Arg426-Gly431 (1291) GGCGCATGTT TGGGGTTCCTGGGGGCC Lys121-Val200-Asn425-Lys432 (1279) GGCGCATGTT TGGGGTTCCTGGGGGCC Val120-Ile201-Ile424-Ala433 (1267) GGCGCATGTT TGGGGTTCCTGGGGGCC (1321) GGCGCATGTT TGGGGTTCCTGGGGGCC (1321) GGCGCCATGTTCCTGGGCGCC (1321) GCCGCCAGCACATGTCGCCCCCAGCACATGTCCTGGCGCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCCAGCACATGTCCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCCAGCACATGTCCCCCCCAGCACATGTCCCCCCCAGCACATGTCCCCCCCAGCACATGTCCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCCAGCACATGTCCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCCAGCACATGTCCCCCCCAGCACATGTCCCCCCCAGCACATGTCCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCAGCACATGTCCCCCCAGCACAACATGCCCCCCAGCACAACAACAACAACAACAACAACAACAACA	Leu122-Ser199-Arg426-Lys432	(1291)	GGGGCATCTICCTGGGCFTGCTGGGGGG
Val120-Ile201-Ile424-Ala433 (1267) GCCCCATCIL TEGGCTTCCTGGCCGCC Val120-Ile201B-Ile424-Ala433 (1267) GCCCCATCIL TEGGCTTCCTGGCCGCC Consensus (1321) GGCGCCATGTTCCTGGGCTTCCTGGGCGCC 1351 1380 Leul22-Ser199 Tryp427-Gly431 (1321) GCCGGCAGCACCATGGCGCGCCGCAGCAGCACCATGGCGCGCCAGCAGCACCATGGCGCCCAGCAGCACCATGGCGCCCCAGCAGCACCATGGCGCCCCCAGCAGCACCATGGCGCCCCCCAGCAGCACCATGGCGCCCCCCCC	Leu122-Ser199-Arg426-Gly431	(1291)	
Val120-Ile201B-Ile424-Ala433 (1267) GCCCATCITC INCGCT CTGGCCGCC Consensus (1321) GGCGCCATGTTCCTGGGCTTCCTGGGCGCC 1351 1380 Leul22-Ser199 Tryp427-Gly431 (1321) GCCGCAGCACCACTGTGCGCGCCCAGCTG Val127-Asn195-Arg426-Gly431 (1351) GCCGCAGCACCATGGGCGCCCCAGCTG Val120-Thr202-Ile424-Ala433 (1297) GCCGCAGCACCATGGGCGCCCCCCAGCTG Leul22-Ser199-Arg426-Lys432 (1321) GCCGCAGCACCATGGGCCCCCCCCAGCTG		(1279)	GGCGCCATGTTCCTGGGGTTCCTGGGCGCC
Val120-Ile2018-Ile424-Ala433 (1267) GSCGCATGITCCTGGGCTGCCTGGGCGCC (1321) GGCGCCATGITCCTGGGCTTCCTGGGCGCC (1351) GGCGCCATGITCCTGGGCTTCCTGGGCGCC (1351) 1380 Leul22-Ser199 Tryp427-Gly431 (1321) GCCGGCAGCACCATGGGCGCCCAGCCTG Val127-Asn195-Arg426-Gly431 (1351) GCCGGCAGCACCATGGGCGCCCCCAGCCTG Val120-Thr202-Ile424-Ala433 (1297) GCCGCAGCACCATGGGCGCCCCCCAGCCTG Leul22-Ser199-Arg426-Lys432 (1321) GCCGCAGCACCATGGGCGCCCCCAGCCTG		(1267)	
Consensus (1321) GGCGCCATGTTCCTGGGCTTCCTGGGCGCC 1351 1380 Leu122-Ser199 Tryp427-Gly431 (1321) GCCGGCAGCACATGGGGGGCGCCAGCATG Val127-Asn195-Arg426-Gly431 (1351) GCCGGCAGCACATGGGGGCCCCCAGCATG Val120-Thr202-Ile424-Ala433 (1297) GCCGCAGCACATGGGGCCCCCCAGCATG Leu122-Ser199-Arg426-Lys432 (1321) GCCGCAGCACATGGGCCCCCCAGCCTG	Vall20-Ile201B-Ile424-Ala433	(1267)	
Leul22-Ser199 Tryp427-Gly431 (1321) GCCGGCAGCACCATGGGGGGCCCCCCAGCTG Val127-Asn195-Arg426-Gly431 (1351) GCCGGCAGCACCATGGGGGCCCCCAGCTG Val120-Thr202-Ile424-Ala433 (1297) GCCGGCAGCACCATGGGGGCCCCGCAGCTG Leul22-Ser199-Arg426-Lys432 (1321) GCCGGCAGCACCATGGGGGCCCCGCAGCTG	Consensus	(1321)	
Vall27-Asn195-Arg426-Gly431 (1351) GCCGCAGCACCATGGGCGCCGCAGCTG Vall20-Thr202-Ile424-Ala433 (1297) GCCGCCAGCACCATGGGCGCCCCCAGCTTG Leul22-Ser199-Arg426-Lys432 (1321) GCCGCCAGCACCACCCCCGCGCCCCCGCCCCCCCCCCC			
Vall27-Asn195-Arg426-Gly431 (1351) GCCGCAGCACCATGGGCGCCGCAGCTG Vall20-Thr202-Ile424-Ala433 (1297) GCCGCCAGCACCATGGGCGCCCCCAGCTTG Leul22-Ser199-Arg426-Lys432 (1321) GCCGCCAGCACCACCCCCGCGCCCCCGCCCCCCCCCCC		(1321)	GCCGGCAGCACCATGGGCGGCCGCAGCCTG
Leul22-Ser199-Arg426-Lys432 (1321) GCC48CAGCACCATGGGCGCCGCAGCCTG	Val127-Asn195-Arg426-Gly431	(1351)	GCCGGCAGCACCATGGGCGCCCGCAGCCTG
Leul22-Ser199-Arg426-Lys432 (1321) GCCGGCAGCACATGGGCGGCGCGCAGCCTG		(1297)	GCCGGCAGCACCATGGGCCCCCCCCCAGCCTG
		(1321)	GCCGCCAGCACCATGGGCGCCCGCAGCCTG
	Leu122-Ser199-Arg426-Gly431	(1321)	

WO 00/39303	36 /	65 PCT/US99/312	72
Lys121-Val200-Asn425-Lys432		GCCGCAGCACCATGGGCGCCCGCAGCCTG	
Val120-Ile201-Ile424-Ala433		GCCGCAGCACCATGGGCGCCCGCAGCCTG	
Val120-Ile201B-Ile424-Ala433	(1297)	GCCGCAGCACCATGGGCGCCCGCAGCCTG	
Consensus	(1351)		
Consensus	(1331;	1381 1410	
Leu122-Ser199 Tryp427-Gly431	(1351)	ACCTGACCGTGCAGGCCCGCCAGCTGCTG	
Vall27-Asn195-Arg426-Gly431	(1381)	The state of the s	
Vall20-Thr202-Ile424-Ala433	(1327)	ACCORGACCGTGCAGGCCCGCCAGCTGCTG	
Leu122-Ser199-Arg426-Lys432	(1351)	ACCCTGACCGTGCAGGCCCGCCAGCTGCTG	
Leu122-Ser199-Arg426-Gly431	(1351)	ACCTGACCGTGCAGGCCCGCCAGCTGCTG	
Lys121-Val200-Asn425-Lys432	(1339)	ACCCTGACCGTGCAGGCCGCCAGCTGCTG	
Val120-Ile201-Ile424-Ala433	(1327)	ACCETGACCGTGCAGGCCGCCAGCTGCTG	
Val120-11e201 11e424-Ala433	(1327)	ACCTGACCGTGCAGGCCGCCAGCTGCTG	
Consensus	(1381)	ACCCTGACCGTGCAGGCCCGCCAGCTGCTG	
consensus	(1301)	1411 1440	
Leu122-Ser199 Tryp427-Gly431	(1381)	AGCGGCATCGTGCAGCAGCAGCAACCTG	
Val127-Asn195-Arg426-Gly431	(1411)	AGCGGCATCGTGCAGCAGCAGCAACCTG	
Val120-Thr202-Ile424-Ala433	(1357)	AGCGGCATCGTGCAGCAGCAGCAACCTG	
Leu122-Ser199-Arg426-Lys432	(1381)	AGCGGCATCGTGCAGCAGCAGCACCACCTG	
Leu122-Ser199-Arg426-Gly431	(1381)	AGCGGCATCGTGCAGCAGCAGCAACCTG	
Lys121-Val200-Asn425-Lys432	(1369)	AGCGGCATCGTGCAGCAGCAACAACCTG	
Val120-Ile201-Ile424-Ala433	(1357)	AGCGGCATCGTGCAGCAGCAGCAACCACCTG	
Val120-Ile201B-Ile424-Ala433	(1357)	AGCGGCATCGTGCAGCAGCAGCAACCTG	
Consensus	(1411)	AGCGGCATCGTGCAGCAGCAGAACAACCTG	
		1441 1470	
Leu122-Ser199 Tryp427-Gly431	(1411)	CTGCGCGCATCGAGGCCCAGCAGCACCTG	
Val127-Asn195-Arg426-Gly431	(1441)	CTGCGCGCATCGAGGCCCAGCAGCACCTG	
Val120-Thr202-Ile424-Ala433	(1387)	CTGCGCGCCATCGAGGCCCCAGCAGCACCTG	
Leu122-Ser199-Arg426-Lys432	(1411)	CTGCGCGCCATCGAGGCCCCAGCAGCACCTG	
Leu122-Ser199-Arg426-Gly431	(1411)	CTGCGCGCCATEGAGGCCCAGCAGCACCTG	
Lys121-Va1200-Asn425-Lys432	(1399)	CTGCGCGCCATCGAGGCCCAGCAGCACCTG	
Val120-Ile201-Ile424-Ala433	(1387)	CTGCGCGCCATGGAGGCCCAGCAGCACCTG	
Val120-Ile201B-Ile424-Ala433	(1387)	CTGGGGGCATCGAGGCCCAGCAGCACCTG	
Consensus	(1441)	CTGCGCGCCATCGAGGCCCAGCAGCACCTG	
		1471 1500	
Leu122-Ser199 Tryp427-Gly431	(1441)	ETGCAGCTGACCGTGTGGGGGATCAAGCAG	
Val127-Asn195-Arg426-Gly431	(1471)	ergeagereacectgtgggggareaageag	
Val120-Thr202-Ile424-Ala433	(1417)	CTGCAGCTGAGCGTGTGGGGCATCAAGCAG	
Leu122-Ser199-Arg426-Lys432	(1441)	ETGCAGCTCACCGTGTGGGGCATCAAGCAG	
Leu122-Ser199-Arg426-Gly431	(1441)	CTGCAGCTGACCGTGTGGGGCATCAAGCAG	
Lys121-Val200-Asn425-Lys432	(1429)	CTGCAGCTGACCGTGTGGGGCATCAAGCAC	
Val120-Ile201-Ile424-Ala433 Val120-Ile201B-Ile424-Ala433	(1417) (1417)	<u>ETGCAGCTGACCGTGTGGGGCATCAAGCAG</u> <u>ETGCAGCTGACCGTGTGGGGCATCAAGCA</u> G	
Consensus	(14/1)	CTGCAGCTGACCGTGTGGGGCATCAAGCAG 1501 1530	
Leu122-Ser199 Tryp427-Gly431	(1471)	enteaveccesecesses sedecesses 1330	
Val127-Asn195-Arg426-Gly431		CTSCAGGCCCCCGTGTGGGCCTGGAGCCC	
Val120-Thr202-Ile424-Ala433		CTSCAGGCCCGCGTGCTGGCCGTCGAGCGC	
Leu122-Ser199-Arg426-Lys432		CTGCAGGCCGCGTGCTGGCCGTGGAGGCC	
Leu122-Ser199-Arg426-Gly431		CTGCAGGCCGCGTGCTGGCGTGGAGGC	
Lys121-Val200-Asn425-Lys432		CTGCAGGGCCGCGTGCTGGCGTGGAGGGC	
Val120-Ile201-Ile424-Ala433		CTGCAGGCCCGCGTGCTGGCGTGGAGCGC	
Val120-Ile201B-Ile424-Ala433		CTGCAGGCCGGCGTGCTGGCCGTGGAGCGC	
Consensus		CTGCAGGCCGCGTGCTGGCCGTGGAGCGC	
2323.000	,	1531 1560	
Leu122-Ser199 Tryp427-Gly431	(1501)	TACCTGAAGGACCAGCAGCTGCTGGGCATC	
Val127-Asn195-Arg426-Gly431	(1531)	TACCTGAAGGACCAGCAGCTGCTGGGCATC	
		The state of the s	

Val120-Thr202-Ile424-Ala433	(1477)	TACCTGAAGGACCAGCTGCTGGGCATC
Leu122-Ser199-Arg426-Lys432	(1501)	TACCTGAAGGACCAGCAGCTGCTGGGCATC
Leu122-Ser199-Arg426-Gly431	(1501)	TACCTGAAGGACCAGCAGCTGCTGGGCATC
Lys121-Val200-Asn425-Lys432	(1489)	NAME (CNACCAMENTACION CONCENCIONIC
Val120-Ile201-Ile424-Ala433	-	
	(1477)	TACCTGARGGACCAGGAGCTGCTGGGCATC
Vall20-Ile201B-Ile424-Ala433	(1477)	TASCIGAAGGAGGAGDAGGTGGTGGGGATG
Consensus	(1531)	TACCTGAAGGACCAGCAGCTGCTGGGCATC
•		1561 1590
Leu122-Ser199 Tryp427-Gly431	(1531)	TGGGGCTGCAGGGGGFAGCTGATCTGCACC
Val127-Asn195-Arg426-Gly431	(1561)	TGGGGCTBCAGCGGGAAGCTGATCTGCACC
Val120-Thr202-Ile424-Ala433	(1507)	TG65GCTGCAGEGBGAAGCTGATCTGCACC
Leu122-Ser199-Arg426-Lys432	(1531)	TOGGGCTGCAGCGCAAGCTGATGTGCACC
-		
Leu122-Ser199-Arg426-Gly431	(1531)	TGGGGCTGCAGCGGCAAGCTGATCTGCACC
Lys121-Val200-Asn425-Lys432	(1519)	TGGGCCTGCAGCGGCAAGCTGATCTGCACC
Val120-Ile201-Ile424-Ala433	(1507)	TGGGGCTGCAGCGGCAGCTGATCTGCACC
Val120-Ile201B-Ile424-Ala433	(1507)	TGGGGCTGCAGCGGCAAGCTGATCTGCACC
Consensus	(1561)	TGGGGCTGCAGCGGCAAGCTGATCTGCACC
		1591 1620
Leu122-Ser199 Tryp427-Gly431	(1561)	ACCGCCGTGCCCTGGAACGCCAGETGGAGC
Val127-Asn195-Arg426-Gly431	(1591)	ACCGCCGTGCCCTGGAACGCCAGCTGGAGC
Val120-Thr202-Ile424-Ala433	(1537)	ACCGCGGTGCCCTGGAACGCCAGGTGGAGC
Leu122-Ser199-Arg426-Lys432	(1561)	ACCGCCGTGCCCTGGAACGCCAGCTGGAGC
Leu122-Ser199-Arg426-Gly431	(1561)	ACCCCCGTGCCCTGGAACGCCAGCTGGAGC
Lys121-Val200-Asn425-Lys432	(1549)	ACCGCGGTGCCCTGGAACGCCAGETGGAGC
Val120-Ile201-Ile424-Ala433	(1537)	ACCGCCGTGCCCTGGAACGCCAGCTGGAGC
Val120-Ile201B-Ile424-Ala433	(1537)	ACCGCCGTGCCCTGGAACGCCAGCTGGAGC
Consensus	(1591)	ACCGCCGTGCCCTGGAACGCCAGCTGGAGC
00113011343	(1331)	1621 1650
100122 Cow100 Www.m427 Cl.4421	(1501)	AACAAGAGCCTGGACEAGATCTGGAACAAC
Leu122-Ser199 Tryp427-Gly431	(1591)	The state of the s
Val127-Asn195-Arg426-Gly431	(1621)	AACAAGAGCCIGGACCAGAICIGGAACAAC
Val120-Thr202-Ile424-Ala433	(1567)	AACAAGAGCCTGGACCAGATCTGGAACAAC
Leu122-Ser199-Arg426-Lys432	(1591)	AACAAGAGCCTEGACCAGATCTGGAACAAC
Leu122-Ser199-Arg426-Gly431	(1591)	AACAAGAGCCTGGACCAGATCTGGAACAAC
Lys121-Val200-Asn425-Lys432	(1579)	AACAAGAGCCTGGACCAGATCTGGAACAAC
Val120-Ile201-Ile424-Ala433	(1567)	AACAAGAGCCTGGACCAGATCTGGAACAAC
Vall20-Ile201B-Ile424-Ala433	(1567)	aacaagageetggaceagatetggaacaac
	(1621)	
Consensus	(1021)	AACAAGAGCCTGGACCAGATCTGGAACAAC
		1651 1680
Leul22-Ser199 Tryp427-Gly431	(1621)	ATGACCTGGATGGAGTGGGAGCGCGAGATC
Val127-Asn195-Arg426-Gly431	(1651)	ATGAUCTGGATGGAGTGGGAGEGUGAGATC
Val120-Thr202-Ile424-Ala433	(1597)	ATGACCTGGATGGAGTGGGAGEGCGAGATC
Leu122-Ser199-Arg426-Lys432	(1621)	ATGACCTGGATGGAGTGGGAGCGCCAGATC
Leu122-Ser199-Arg426-Gly431	(1621)	ANGACCTEGATEGAGUEGGACCGCGAGATC
Lys121-Va1200-Asn425-Lys432	(1609)	ATGACCTGCATCCACTCCCACGCGAGATC
Val120-Ile201-Ile424-Ala433	(1597)	AND REVICEABLESACHEEDAGGECAGATE
Val120-Ile201B-Ile424-Ala433	(1597)	AND COLORS AND
		ATGACCTGGATGGAGTGGGAGCGCGAGATC
Consensus	(1651)	ATGACCTGGATGGAGTGGGAGCGCGAGATC
		1681 1710
Leul22-Ser199 Tryp427-Gly431	(1651)	GACAACTACACCAACCTGATCTACACCCTG
Val127-Asn195-Arg426-Gly431	(1681).	GACAACTACACCAACCTGATCTACACCCTG
Val120-Thr202-Ile424-Ala433	(1627)	GACAACTAGACCAACCTGATGTACACCCTG
Leu122-Ser199-Arg426-Lys432	(1651)	GACANCTACACCAACCTGATCTACACCCTG
Leu122-Ser199-Arg426-Gly431	(1651)	GACAACTACACCAACCTGATCTACACCCTG
Lys121-Val200-Asn425-Lys432	(1639)	
•		GACAACTACACCAACCTGATCTACACCCTG
Val120-Ile201-Ile424-Ala433	(1627)	GACAACTACACCAACTGATCTACACCCTG
Vall20-Ile201B-Ile424-Ala433	(1627)	GACAACTACACCAACCTGATCTACACCCTG
Consensus	(1681)	GACAACTACACCAACCTGATCTACACCCTG

WO 00/39303	38	/ 65	PCT/US99/31272
		1711	1740
Leu122-Ser199 Tryp427-Gly431	(1681)		AGCCAGAACCAGCAGAAG
Val127-Asn195-Arg426-Gly431	(1711)		AGCCAGAACCAGCAGGAGAAG
Val120-Thr202-Ile424-Ala433	(1657)		AGCCAGAACCAGCAGGAGAAG
Leu122-Ser199-Arg426-Lys432	(1681)		AGCCAGAACCAGCAGGAGAAG
Leu122-Ser199-Arg426-Gly431	(1681)		AGCCAGAACCAGCAGAGAAG
Lys121-Val200-Asn425-Lys432	(1669)		AGECAGAACCAGCAGAAAG
Val120-Ile201-Ile424-Ala433 Val120-Ile201B-Ile424-Ala433	(1657)	· chileran and annual contract	AGCCAGAACCAGCAGGAGAAG AGCCAGAACCAGCAGGAGAAG
Consensus	(1657) (1711)		AGCCAGAACCAGCAGGAGAAG
Consensus	(1/11)	1741	1770
Leu122-Ser199 Tryp427-Gly431	(1711)		GAGCTGCTGGAGCTGGACAAG
Val127-Asn195-Arg426-Gly431	(1741)		GAGCTGCTGGAGCAAG
Val120-Thr202-Ile424-Ala433	(1687)		<u>GAGCTGCTGGAGCAAG</u>
Leu122-Ser199-Arg426-Lys432	(1711)		GÁGCTGCTGGAGCAAG
Leu122-Ser199-Arg426-Gly431	(1711)		GAGCTGCTGGAGCTGGACAAG
Lys121-Val200-Asn425-Lys432	(1699)		GAGCTGCTGGAGCAAG
Val120-Ile201-Ile424-Ala433	(1687)		GAGCTGCTGGAGCAAG
Val120-Ile201B-Ile424-Ala433	(1687)		GAGCTGCTGGAGCTGGACAAG
Consensus	(1741)		GAGCTGCTGGAGCTGGACAAG
		1771	1800
Leu122-Ser199 Tryp427-Gly431	(1741)	TGGGCCAGC	CTGTGGAACTGGTTCGAGATC
Val127-Asn195-Arg426-Gly431	(1771)		<u>CTGTGGAACTGGTTCGACATC</u>
Val120-Thr202-Ile424-Ala433	(1717)		<u>ETGTGGAACTGGTTCGACATC</u>
Leu122-Ser199-Arg426-Lys432	(1741)		CTGTGGAACTGGTTCGACATC
Leu122-Ser199-Arg426-Gly431	(1741)		CTGTGGAACTGGTTCGACATC
Lys121-Val200-Asn425-Lys432	(1729)		CTGTGGAACTGGTTCGAGATC
Vall20-Ile201-Ile424-Ala433	(1717)		CTGTGGAACTGGTTCGACATC
Val120-Ile201B-Ile424-Ala433	(1717)		CTGTGGAACTGGTTCGACATC
Consensus	(1771)	TGGGCCAGC	CTGTGGAACTGGTTCGACATC 1830
Leu122-Ser199 Tryp427-Gly431	(1771)		CTGTGGTACATCAAGATCTTC
Val127-Asn195-Arg426-Gly431	(1801)		CAGAGGAAGAAGAAGAA
Val120-Thr202-Ile424-Ala433	(1747)		GREFEGIACATCANGATE TO
Leu122-Ser199-Arg426-Lys432	(1771)		CIGIGGTACATCAAGATCITC
Leu122-Ser199-Arg426-Gly431	(1771)		CTGTGGTACATCAAGATCTTC
Lys121-Val200-Asn425-Lys432	(1759)	AGCAAGTGG	CIGIGGTACATCAAGATCTTC
Val120-Ile201-Ile424-Ala433	(1747)		CTGTGGTACATCAAGATCTTC
Val120-Ile201B-Ile424-Ala433	(1747)		CTGTGGTACATCAAGATCTTC
Consensus	(1801)		CTGTGGTACATCAAGATCTTC
		1831	1860
Leu122-Ser199 Tryp427-Gly431	(1801)		GTGGGGGGCCTGGTGGGGGTG
Val127-Asn195-Arg426-Gly431	(1831)		GTGGGCGGCCTGGTGGGCCTG
Val120-Thr202-Ile424-Ala433		ATCATGATC	<u>creesegectestescet</u> e
Leu122-Ser199-Arg426-Lys432	(1801)		eregececteteteecere
Leu122-Ser199-Arg426-Gly431	(1801)		GREECECCREGREECE RE
Lys121-Val200-Asn425-Lys432	(1789)		eregececteetegge <u>r</u> e
Val120-Ile201-Ile424-Ala433	(1777)		GTGGGCGCCTGGTGGGCCTG
Vall20-Ile201B-Ile424-Ala433	(1777)		STGGGCGGCCTGGTGGGCCTG
Consensus	(1831)	1861	GTGGGCGGCCTGGTGGGCCTG 1890
Leu122-Ser199 Tryp427-Gly431	(1831)		TTCACCGTGCTGAGCATCGTG
Val127-Asn195-Arg426-Gly431	(1861)		TTCACCGTGCTGAGCATCGTG
Val120-Thr202-Ile424-Ala433	(1807)		TTCACCGTGCTGAGCATCGTG
Leul22-Ser199-Arg426-Lys432	(1831)		LICACCGTGCTGAGCATCGTG
Leu122-Ser199-Arg426-Gly431	(1831)	CGCATCGTG	ITCACCGTGCTGAGCATCGTG
Lys121-Val200-Asn425-Lys432	(1819)		TTCACCGTGCTGAGCATCGTG

Val120-Ile201-Ile424-Ala433	(1807)	
Vall20-Ile201B-Ile424-Ala433	(1807)	
Consensus	(1861)	
Leul22-Ser199 Tryp427-Gly431	110611	1891 1920
Vall27-Asn195-Arg426-Gly431	(1861)	
Val120-Thr202-Ile424-Ala433	(1891)	
Leu122-Ser199-Arg426-Lys432	(1837) (1861)	
Leu122-Ser199-Arg426-Gly431	(1861)	
Lys121-Val200-Asn425-Lys432	(1849)	
Val120-Ile201-Ile424-Ala433	(1837)	
Val120-Ile201B-Ile424-Ala433	(1837)	
Consensus	(1891)	
	(10)1,	1921 1950
Leu122-Ser199 Tryp427-Gly431	(1891)	
Val127-Asn195-Arg426-Gly431	(1921)	AGCTTCCAGACCCGCTTCCCCGCCCCCCCC
Vall20-Thr202-Ile424-Ala433	(1867)	AGCTTCCAGACCCGCTTCCCCGCCCCCGC
Leu122-Ser199-Arg426-Lys432	(1891)	AGCTTCCAGACECGCTTCCCCGCCCCGCGC
Leu122-Ser199-Arg426-Gly431	(1891)	AGETTECAGAGECGETTEGECGCCCCCCCCC
Lys121-Val200-Asn425-Lys432	(1879)	AGCTTCCAGACCCGCTTCCCCCCCCCCCC
Val120-Ile201-Ile424-Ala433	(1867)	AGETTCCAGACCCGCTTCCGCGCCCCCGC
Vall20-Ile201B-Ile424-Ala433	(1867)	AGCTTECAGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
Consensus	(1921)	AGCTTCCAGACCCGCTTCCCCGCCCCCGC
Y 100 0 100 m		1951 1980
Leu122-Ser199 Tryp427-Gly431	(1921)	GGCCCGACCGCCCGAGGGCATCGAGGAG
Vall27-Asn195-Arg426-Gly431	(1951)	GGCCCCGACGCCCCGAGGCATCGAGGAG
Val120-Thr202-Ile424-Ala433	(1897)	GGCCCGAGCGCCCGAGGGCATCGAGGAG
Leu122-Ser199-Arg426-Lys432	(1921)	GGCCCGACCGCCCCGAGGGCATCGAGGAG
Leu122-Ser199-Arg426-Gly431	(1921)	GGCCCCGACCCCCCCGACCCCACCCACCCAC
Lys121-Val200-Asn425-Lys432 Val120-Ile201-Ile424-Ala433	(1909)	acoustances and lakes in a latter of the latter of
Val120-11e201-11e424-A1a433 Val120-11e201B-11e424-A1a433	(1897)	ପ୍ରକ୍ରେକ୍ଟ: ୧୭୮୯: ବ୍ୟକ୍ତ: (୧୯୮୯: ୬/୩୯୯୮) ମଧ୍ୟ
	(1897)	ECOCOCCACCOSOCCA GESCOATICGA CEAG
Consensus	(1951)	GGCCCGACCGCCCGAGGGCATCGAGGAG 1981 2010
Leu122-Ser199 Tryp427-Gly431	(1951)	1981 2010 GAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
Val127-Asn195-Arg426-Gly431	(1981)	GAGGGCGGCGAGCCGCAGC
Val120-Thr202-Ile424-Ala433	(1927)	GAGGGGGGGAGGGGAGGGGAGG
Leu122-Ser199-Arg426-Lys432	(1951)	CAGGGGGGGAGGGGAGGGGAGG
Leu122-Ser199-Arg426-Gly431	(1951)	GNedendend valendy dated verse nec
Lys121-Val200-Asn425-Lys432	(1939)	GAGGGGGGGAGGGGAGGGGAGG
Val120-Ile201-Ile424-Ala433	(1927)	GAGGGGGGGAGAGGGAGAGGGAAGGGGAAGG
Val120-Ile201B-Ile424-Ala433	(1927)	ପ୍ର:(୯୯୯,୧ଟମ୍ବର୍ଷ ,୯୯୭୯୫୯,୧୭୯,୧୯୭୯,୧୯
Consensus	(1981)	GAGGGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
		2011 2040
Leu122-Ser199 Tryp427-Gly431	(1981)	YCL-acom recine veccement en recionarie
Val127-Asn195-Arg426-Gly431	(2011)	AGCCCCCTICETE OF A GCC OF ITE CITE OF CITE
Vall20-Thr202-Ile424-Ala433	(1957)	AGCCCCTGGTGCACGGCCTGCTGGCCCTG
Leu122-Ser199-Arg426-Lys432	(1981)	AGCCCCCTGGTGCACGGCCTGCTGGCCCTG
Leu122-Ser199-Arg426-Gly431	(1981)	AGCCCCTEGTECACEGCCTGGTGCCCCTG
Lys121-Val200-Asn425-Lys432	(1969)	AGCECCTGGTGCACGGCCTGCTGGCCCTG
Val120-Ile201-Ile424-Ala433	(1957)	AGCCCCTGGTGCACGGCCTGCTGGCCCTG
Vall20-Ile201B-Ile424-Ala433	(1957)	AGCCCCCTGGTGCACGGCCTGCTGGCCCTG
Consensus		AGCCCCTGGTGCACGGCCTGCTGGCCCTG
Leul 22-Sow100 m 407 01 101		2041 2070
Leul22-Ser199 Tryp427-Gly431 Val127-Asn195-Arg426-Gly431	(2011)	VARIANCE COVICE SOCIAL CONTRACTOR
Vall20-Thr202-Ile424-Ala433	(2041)	atcteggacgacetgcegaccetgtgeetg
INITEDS-116454-W19433	(1987)	atcheegaceasetecetere

```
(2011) ATCTGGGACGACCTGCGCAGCCTGTGGCTG
   Leu122-Ser199-Arg426-Lys432
   Leu122-Ser199-Arg426-Gly431
                                                    (2011) ATCTGGGACGACCTGCGCAGCCTGTGCCTG
   Lys121-Val200-Asn425-Lys432
                                                    (1999) ATCTGGGACGACCTGCGCAGCCTGTGCCTG
   Val120-Ile201-Ile424-Ala433
                                                   (1987) ATCTGGGACGACCTGCGCAGCCTGTGCCTG
                                                   (1987) ATCTGGGACGACCTGCGCAGCCTGTGCCTG
  Val120-Ile201B-Ile424-Ala433
                                 Consensus
                                                   (2041) ATCTGGGACGACCTGCGCAGCCTGTGCCTG
                                                               2071
 Leu122-Ser199 Tryp427-Gly431
                                                   (2041) TTCAGCTACCACCGCCTGCGCGACCTGATC
   Val127-Asn195-Arg426-Gly431
                                                   (2071) TTCAGCTACCACCGCCTGCGGGACCTGATC
   Val120-Thr202-Ile424-Ala433
                                                   (2017) TTCAGCTACCACCGCCTGCGCGACCTGATC
   Leu122-Ser199-Arg426-Lys432
                                                   (2041) TICAGCTACCACCGCCTGCGCGACCTGATC
   Leu122-Ser199-Arg426-Gly431
                                                   (2041) TTCAGCTACCACCGCCTGCGCGCGACCTGATC
   Lys121-Val200-Asn425-Lys432
                                                   (2029) TICAGCTACCACCGCCTGCGCGACCTGATC
   Val120-Ile201-Ile424-Ala433
                                                   (2017) TTCAGCTACCACCGCCTGCGCGACCTGATC
 Val120-Ile201B-Ile424-Ala433
                                                   (2017) TTCAGCTACCACCGCCTGCGCGACCTGATC
                                Consensus
                                                   (2071) TTCAGCTACCACCGCCTGCGCGACCTGATC
 Leu122-Ser199 Tryp427-Gly431
                                                   (2071) CTGATEGECGCCCGCATEGTGGAGCTGCTG
   Val127-Asn195-Arg426-Gly431
                                                   (2101) CTGATCGCCGCCCGCATCGTGGAGCTGCTG
   Val120-Thr202-Ile424-Ala433
                                                   (2047)
                                                             CTGATEGCCGCCCCCATCGTGGAGCTGCTG
  Leu122-Ser199-Arg426-Lys432
                                                   (2071) CTGATEGEGGCCGCATCGTGGAGCTGCTG
  Leu122-Ser199-Arg426-Gly431
                                                   (2071) CTGATGGCCGCCGCATCGTGCAGCTGCTG
  Lys121-Val200-Asn425-Lys432
                                                   (2059) CTGATEGECGCCCGCATEGTGGAGCTGETG
   Val120-Ile201-Ile424-Ala433
                                                  (2047) CTGATEGECGCCCGCATEGTGGAGCTGCTG
 Val120-Ile201B-Ile424-Ala433
                                                   (2047) CTGATCGCCGCCCGCATCGTGGAGCTGCTG
                                Consensus
                                                  (2101) CTGATCGCCGCCCGCATCGTGGAGCTGCTG
                                                              2131
                                                                                                         2160
Leu122-Ser199 Tryp427-Gly431
                                                   (2101) GGCCGCCGCGCTGGGGCCCTGAGTAC
  Val127-Asn195-Arg426-Glv431
                                                  (2131) GGCCGCCGCGCTGGGAGCCCCTGAAGTAC
                                                  (2077) GGCCGCGGGGGCTGGGAGCCCTGAAGTAC
  Val120-Thr202-Ile424-Ala433
  Leu122-Ser199-Arg426-Lys432
                                                  (2101) GGCCGCGCGGGGGGGGCCCTGAAGTAC
                                                             GGGGGHVCRGGALGCGYGEGGGAGYGYYGYY
GGGGGGGGGGGGGAGGAGGAGGAGA
  Leu122-Ser199-Arg426-Gly431
                                                  (2101)
  Lys121-Val200-Asn425-Lys432
                                                  (2089)
                                                  Val120-Ile201-Ile424-Ala433
Val120-Ile201B-Ile424-Ala433
                                                  (2077)
                                                             GGCCGCGGGGTGGGGGGGCCCTGAAGTAC
                                                  (2131) GGCCGCCGCGGCTGGGAGGCCCTGAAGTAC
                               Consensus
                                                             2161
Leu122-Ser199 Tryp427-Gly431
                                                  (2131)
                                                             REGEROAL CONFERNICATION OF THE PROPERTY OF THE
  Val127-Asn195-Arg426-Gly431
                                                  (2161)
                                                             TGGGGCAACCTGCTGCAGTACTGGATCCAG
  Val120-Thr202-Ile424-Ala433
                                                  (2107)
                                                             TGGGGGACCTGCTGCAGTACTGGATCCAG
  Leu122-Ser199-Arg426-Lvs432
                                                  (2131) TGGGG: AAUGTGGTGCAGTAGTGGATCCAG
 Leu122-Ser199-Arg426-Gly431
                                                  (2131) TGGGGBAACCTGCTGCAGPACTCCATCCAG
 Lys121-Val200-Asn425-Lys432
                                                             TGGGGGAAGCTGCTGCAGTAGTAGTAGTCCAG
                                                  (2119)
 Val120-Ile201-Ile424-Ala433
                                                  (2107)
                                                             HELECTIVE CANENT COVERNER CONTROL (C
Val120-Ile201B-Ile424-Ala433
                                                  (2107)
                                                             TGGGGCAACCTGCTGCAGTACTGGATCCAG
                               Consensus
                                                  (2161) TGGGGCAACCTGCTGCAGTACTGGATCCAG
                                                                                                        2220
Leu122-Ser199 Tryp427-Gly431
                                                  (2161) GAGGTGAAGAACAGCGCCGTGAGGGTGATG
 Val127-Asn195-Arg426-Gly431
                                                  (2191)
                                                             GAGCTGAAGAACAGCGCCGTGAGCCTGTTC
 Val120-Thr202-Ile424-Ala433
                                                             GAGCTGAAGAACAGCGCCGTGAGCCTGTTC
                                                  (2137)
 Leul22-Ser199-Arg426-Lys432
                                                  (2161)
                                                             GAGCTGAAGAACAGCGCCGTGAGCCTGTTC
 Leu122-Ser199-Arg426-Gly431
                                                  (2161)
                                                             GAGCTGAGCACAGCGCCGTGAGCCTGTTC
 Lys121-Val200-Asn425-Lys432
                                                  (2149) GAGCTGAAGAACAGCGCCGTGAGCCTGTTC
 Val120-Ile201-Ile424-Ala433
                                                  (2137) GAGCTGAAGAACAGCGCCGTGAGCCTGTTC
Val120-Ile201B-Ile424-Ala433
                                                  (2137) GAGCTGAAGAACAGCGCCGTGAGCCTGTTC
                               Consensus
                                                  (2191) GAGCTGAAGAACAGCGCCGTGAGCCTGTTC
```

SEQ ID NO:3 VAL120-ALA204

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCTGCGTGCCCACCGACCCCAACCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGGGCGCCGGCGCCTGCCCCAA GGTGAGCTTCGAGCCCATCCCATCCACTACTGCGCCCCGGCTTCGCCATCCTGAAGTG CAACGACAAGAAGTTCAACGGCAGCGGCCCCTGCACCAACGTGAGCACCGTGCAGTGCACCC ACGGCATCCGCCCGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCCGAGGAGGGC GTGGTGATCCGCAGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCTGAAGGA GAGCGTGGAGATCAACTGCACCCGCCCAACAACAACACCCGCAAGAGCATCACCATCGGCC CCGGCCGCCTTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACA TCAGCGGCGAGAAGTGGAACACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAGTTC GGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGATGCACAG CTTCAACTGCGGCGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAA CAACACCATCGGCCCCAACAACACCAACGGCACCATCACCCTGCCCTGCCGCATCAAGCAGA TCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATGTACGCCCCCCCATCCGCGGCCAGATC CGCTGCAGCAGCAACATCACCGGCCTGCTGCTGACCCGCGACGGCGCAAGGAGATCAGCAA CACCACCGAGATCTTCCGCCCCGGCGCGCGCGACATGCGCGACAACTGGCGCAGCGAGCTGT ACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCACCAAGGCCAAGCGCCGC GTGGTGCAGCGCGAGAAGCGCGCCGTGACCCTGGGCGCCATGTTCCTGGGCTTCCTGGGCGCC GCCGCCACCATGGCCCCGCAGCCTGACCCTGCCGCCAGCCGCCAGCTGCTGAG CGGCATCGTGCAGCAGCAGCACCACCTGCTGCGCGCCCATCGAGGCCCAGCAGCACCTGCTGC AGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTG AAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGT GCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAGATCTGGAACAACATGACCTGGA TGGAGTGGGAGCGCGAGATCGACAACTACACCAACCTGATCTACACCCTGATCGAGGAGAGC CAGAACCAGCAGGAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCAGCCTGT GGAACTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCG GCCTGGTGGGCCTGCGCATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCT ACAGCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCGCGCCCCGACCGCCCCGAGGGCA TCGAGGAGGAGGCGCGACCGCGACCGCACCGCAGCCCCCTGGTGCACGGCCTGCTG GCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCAGCTACCACCGCCTGCGCGACCTG ATCCTGATCGCCGCCCGCATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGGCCCTGAAGTAC TGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTCGA CGCCATCGCCATCGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCCCAGCGCATCG GCCGCGCCTTCCTGCACATCCCCCGCCGCATCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAAC **TCGAG**

SEQ ID NO:4 VAL120-ILE201

GAATTCGCCACCATGGATGCAATGAAGAGGGCTCTGCTGTTGTTGCTGTTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGCCACCACCCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGGGCGGCATCACCCAGGCCTG CCCCAAGGTGAGCTTCGAGCCCATCCCCATCCACTACTGCGCCCCGCCGGCTTCGCCATCCT GAAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCCTGCACCAACGTGAGCACCGTGCAGT GCACCCACGGCATCCGCCCGTGGTGAGCACCCAGCTGCTGCAACGGCAGCCTGGCCGAG GAGGGCGTGGTGATCCGCAGCGAGACTTCACCGACAACGCCAAGACCATCATCGTGCAGCT GAAGGAGAGCGTGGAGATCAACTGCACCCGCCCCAACAACAACACCCGCAAGAGCATCACCA TCGGCCCGGCCGCCTTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACT GCAACATCAGCGGCGAGAAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCC CAGTTCGGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGAT GCACAGCTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCAC CTGGAACAACACCATCGGCCCCAACAACACCCAACGGCACCATCACCCTGCCCTGCCGCATCA CAGATCCGCTGCAGCAGCACATCACCGGCCTGCTGCTGACCCGCGACGGCGCAAGGAGAT CAGCAACACCACGAGATCTTCCGCCCGGCGGCGGCGACATGCGCGACAACTGGCGCAGCG AGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCACCAAGGCCAAG CGCCGCGTGGTGCAGCGCGAGAAGCGCGCCGTGACCCTGGGCGCCATGTTCCTGGGCTTCCTG GGCGCCGCCGCAGCACCATGGGCGCCCGCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCT GCTGAGCGGCATCGTGCAGCAGCAGCACCACCTGCTGCGCCCATCGAGGCCCAGCAGCACC TGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGC TACCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCAC CGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAGATCTGGAACAACATGA CCTGGATGGAGTGGGAGCGCGAGATCGACAACTACACCAACCTGATCTACACCCTGATCGAG GAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGAGCAGAGTGGGCCA GCCTGTGGAACTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCG TGGGCGGCTGGTGGGCCTGCGCATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGTGCGCC AGGGCTACAGCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCCGGGCCCCGACCGCCCCG AGGGCATCGAGGAGGGGGGGGGGGGGGGGGGCGACCGCAGCAGCCCCCTGGTGCACGG CCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCAGCTACCACCGCCTGCG CGACCTGATCCTGATCGCCGCCCGCATCGTGGAGCTGCTGGGCCGCCGCGCTGGGAGGCCCT GAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCC TGTTCGACGCCATCGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCCCAGC GCATCGGCCGCGCTTCCTGCACATCCCCCGCCGCATCCGCCAGGGCTTCGAGCGCGCCCTGC **TGTAACTCGAG**

SEQ ID NO:5 VAL120-ILE201B

TTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCCA CCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGTGCACAACGTGTGGGCCACCC ACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCTGGAGAACGTGACCGAGAACTTCAACA TGTGGAAGAACAACATGGTGGAGCAGATGCACGAGGACATCATCAGCCTGTGGGACCAGAGCCTGAAGC CCTGCGTGCCCGGCATCACCCAGGCCTGCCCCAAGGTGAGCTTCGAGCCCATCCCCATCCACTACTGCGC CCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCAGCGCCCCTGCACCAACGT GAGCACCGTGCAGTGCACCCACGGCATCCGCCCCGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCT GGCCGAGGAGGCCTGGTGATCCGCAGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCT GAAGGAGAGCGTGGAGATCAACTGCACCCGCCCAACAACAACACCCGCAAGAGCATCACCATCGGCCC CGGCCGCGCTTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGC GAGAAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAGTTCGGCAACAAGACCATC GTGTTCAAGCAGAGCAGCGGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCGGCGGCGAGTTC TTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCGGCCCCAACAACACCAAC GGCACCATCACCCTGCCCTGCCGCATCAAGCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATG TACGCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCAACATCACCGGCCTGCTGCTGACCCGCGACG GCGGCAAGGAGATCAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGGCGACATGCGCGACAACTGGC GCAGCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCACCAAGGCCAAGC GCCGCGTGGTGCAGCGCGAGAAGCGCGCCGTGACCCTGGGCGCCATGTTCCTGGGCTTCCTGGGCGCCGC CGGCAGCACCATGGGCGCCCGCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGAGCGGCATCGT GCAGCAGCAGAACAACCTGCTGCGCGCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGG CATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGAGCGCTACCTGAAGGACCAGCAGCTGCTGGGCAT CTGGGGCTGCAGCGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAG CCTGGACCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGCGAGATCGACAACTACACCAACCT GATCTACACCCTGATCGAGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTGG ACAAGTGGGCCAGCCTGTGGAACTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCAT GATCGTGGGCCTGGTGGGCCTGCGCATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGTGCGCCAG GGCTACAGCCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCGGGGCCCCGACGCCCCGAGGGCATCG AGGAGGAGGCGCGAGCGCGACCGCGACCGCAGCCCCCTGGTGCACGGCCTGCTGGCCCTGATCT GGGACGACCTGCGCAGCCTGTGCCTGTTCAGCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCCG CATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGCTGCAGTACTG GATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCATCGCCGTGGCCGAGGGCAC CGACCGCATCATCGAGGTGGCCCAGCGCATCGGCCGCCCTTCCTGCACATCCCCCGCCGCATCCGCCAG GGCTTCGAGCGCGCCCTGCTGTAACTCGAGCGTGCT

SEQ ID NO:6 LYS121-VAL200

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGCCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGGCCCCCGTGATCACCCA GGCCTGCCCAAGGTGAGCTTCGAGCCCATCCCATCACTACTGCGCCCCGCCGGCTTCGC CATCCTGAAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCCTGCACCAACGTGAGCACCG TGCAGTGCACCCACGGCATCCGCCCCGTGGTGAGCACCCAGCTGCTGCAACGGCAGCCTGG CCGAGGAGGCGTGGTGATCCGCAGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTG CAGCTGAAGGAGCGTGGAGATCAACTGCACCCGCCCCAACAACAACACCCGCAAGAGCAT CACCATCGGCCCGGCCGCCTTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGC CCACTGCAACATCAGCGGCGAGAAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGC AGGCCCAGTTCGGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATC GTGATGCACAGCTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAAC AGCACCTGGAACAACACCATCGGCCCCAACAACACCCAACGGCACCATCACCCTGCCCTGCCG CATCAAGCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATGTACGCCCCCCCATCC GCGGCCAGATCCGCTGCAGCAGCAACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAAG GAGATCAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGGCGACATGCGCGACAACTGGCG CAGCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCACCAAGG CCAAGCGCCGCGTGGTGCAGCGCGAGAAGCGCGCCGTGACCCTGGGCGCCATGTTCCTGGGC TTCCTGGGCGCCGGCAGCACCATGGGCGCCCGCAGCCTGACCCTGACCGTGCAGGCCCGC CAGCTGCTGAGCGGCATCGTGCAGCAGCAGAACAACCTGCTGCGCGCCATCGAGGCCCAGCA GCACCTGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGG AGCGCTACCTGAAGGACCAGCAGCTGCTGGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGC ACCACCGCCGTGCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAGATCTGGAACAA CATGACCTGGATGGAGTGGGAGCGCGAGATCGACAACTACACCAACCTGATCTACACCCTGA TCGAGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGACAAGTG GGCCAGCCTGTGGAACTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCAT GATCGTGGGCGGCCTGGTGGGCCTGCGCATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGT GCGCCAGGGCTACAGCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCCGCGGCCCCGACCG CCCCGAGGCATCGAGGAGGAGGGCGCGAGCGCGACCGCACCGCAGCAGCCCCCTGGTGC ACGCCTGCTGGCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCAGCTACCACCGCC TGCGCGACCTGATCCTGATCGCCGCCCGCATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGG CCCTGAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTG CAGCGCATCGGCCGCCCTTCCTGCACATCCCCCGCCGCATCCGCCAGGGCTTCGAGCGCGCC CTGCTGTAACTCGAGCGTGCT

SEQ ID NO:7: LEU122-SER199

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCTGCGTGAAGCTGGGCAACAGCGTGAT CACCCAGGCCTGCCCCAAGGTGAGCTTCGAGCCCATCCCCATCCACTACTGCGCCCCGCCGG CTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCCTGCACCAACGTGA GCACCGTGCAGTGCACCCACGGCATCCGCCCCGTGGTGAGCACCCAGCTGCTGCTGAACGGC AGCCTGGCCGAGGAGGGCGTGGTGATCCGCAGCGAGAACTTCACCGACAACGCCAAGACCAT CATCGTGCAGCTGAAGGAGCGTGGAGATCAACTGCACCCGCCCCAACAACAACACCCGCA AGAGCATCACCATCGGCCCGGCCGCCCTTCTACGCCACCGGCGACATCATCGGCGACATCC GCCAGGCCCACTGCAACATCAGCGGCGAGAAGTGGAACAACACCCTGAAGCAGATCGTGACC AAGCTGCAGGCCCAGTTCGGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGGCGGCGACCC CGAGATCGTGATGCACAGCTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACCCAGCT GTTCAACAGCACCTGGAACAACACCATCGGCCCCAACAACACCAACGGCACCATCACCCTGC CCTGCCGCATCAAGCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATGTACGCCCCC CCCATCCGCGGCCAGATCCGCTGCAGCAGCAACATCACCGGCCTGCTGCTGACCCGCGACGGC GGCAAGGAGATCAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGGCGACATGCGCGACAA CTGGCGCAGCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCA CCAAGGCCAAGCGCGCGTGTGCAGCGCGAGAAGCGCGCGTGACCCTGGGCGCCATGTTC CTGGGCTTCCTGGGCGCCGCGGCAGCACCATGGGCGCCCGCAGCCTGACCCTGACCGTGCAG GCCCGCCAGCTGCTGAGCGGCATCGTGCAGCAGCAGCACCACCTGCTGCGCGCCCATCGAGGC CCAGCAGCACCTGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGG CCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCAGCGGCAAGCTG ATCTGCACCACCGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAGATCTG GAACAACATGACCTGGATGGAGTGGGAGCGCGAGATCGACAACTACACCAACCTGATCTACA CCCTGATCGAGGAGAGCCAGAACCAGCAGGAGAACAGCAGGAGCTGCTGGAGCTGGA CAAGTGGCCAGCCTGTGGAACTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTT CATCATGATCGTGGGCCTGGTGGGCCTGCGCATCGTGTTCACCGTGCTGAGCATCGTGAA CCGCGTGCGCCAGGCTACAGCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCCGCGGCCC CGACCGCCCGAGGGCATCGAGGAGGAGGGCGGCGACCGCGACCGCAGCAGCAGCCCC $\tt CTGGTGCACGGCCTGGCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCAGCTAC$ CACCGCCTGCGCGACCTGATCCTGATCGCCGCCCGCATCGTGGAGCTGCTGGGCCGCCGCGGC TGGGAGGCCCTGAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAACAG GGTGGCCCAGCGCATCGGCCGCCTTCCTGCACATCCCCCGCCGCATCCGCCAGGGCTTCGA GCGCGCCCTGCTGTAACTCGAGCGTGCT

SEO ID NO:8 VAL120-THR202

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGGGCGGCGCCACCCAGGCCTG CCCCAAGGTGAGCTTCGAGCCCATCCCCATCCACTACTGCGCCCCGCCGGCTTCGCCATCCT GAAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCCTGCACCAACGTGAGCACCGTGCAGT GCACCCACGGCATCCGCCCGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCCGAG GAGGGCGTGGTGATCCGCAGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCT GAAGGAGAGCGTGGAGATCAACTGCACCCGCCCCAACAACACCCCGCAAGAGCATCACCA TCGGCCCGGCCGCCTTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACT GCAACATCAGCGGCGAGAAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCC CAGTTCGGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGAT GCACAGCTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCAC CTGGAACAACACCATCGGCCCCAACAACACCAACGGCACCATCACCCTGCCCTGCCGCATCA CAGATCCGCTGCAGCAGCAACATCACCGGCCTGCTGACCCGCGACGGCGGCAAGGAGAT CAGCAACACCACGAGATCTTCCGCCCGGCGGCGGCGACATGCGCGACAACTGGCGCAGCG AGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCACCAAGGCCAAG CGCCGCGTGTGCAGCGCGAGAAGCGCGCCGTGACCCTGGGCGCATGTTCCTGGGCTTCCTG GGCGCCGCGGCACCATGGGCGCCCGCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCT GCTGAGCGGCATCGTGCAGCAGCAGCACCACCTGCTGCGCGCCATCGAGGCCCAGCAGCACC TGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGC TACCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCAC CGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAGATCTGGAACAACATGA CCTGGATGGAGTGGGAGCGCGAGATCGACAACTACACCAACCTGATCTACACCCTGATCGAG GAGAGCCAGAACCAGCAGGAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCA GCCTGTGGAACTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCG TGGGCGGCTGGTGGGCCTGCGCATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGTGCGCC AGGGCTACAGCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCGGGCCCCGACCGCCCCG AGGGCATCGAGGAGGAGGGCGGCGAGCGCGACCGCAGCAGCCCCCTGGTGCACGG CCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCAGCTACCACCGCCTGCG CGACCTGATCCTGATCGCCGCCCGCATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGGCCCT GAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCC TGTTCGACGCCATCGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCCCAGC GCATCGGCCGCCTTCCTGCACATCCCCGCCGCATCCGCCAGGGCTTCGAGCGCGCCCTGC **TGTAACTCGAG**

SEQ ID NO:9 TRP427-GLY431

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCTGCGTGAAGCTGACCCCCCTGTGCGTG ACCCTGCACTGCACCAACCTGAAGAACGCCACCAACACCAAGAGCAGCAACTGGAAGGAGAT GGACCGCGGCGAGATCAAGAACTGCAGCTTCAAGGTGACCACCAGCATCCGCAACAAGATGC AGAAGGAGTACGCCCTGTTCTACAAGCTGGACGTGCCCCATCGACAACGACAACACCAGC TACAAGCTGATCAACTGCAACACCAGCGTGATCACCCAGGCCTGCCCCAAGGTGAGCTTCGA GCCATCCCATCACTACTGCGCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAA GTTCAACGGCAGCGCCCCTGCACCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCC CCGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCCGAGGAGGGCGTGGTGATCCGC AGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGAGAT CAACTGCACCGCCCAACAACACCCGCAAGAGCATCACCATCGGCCCCGGCCGCCCT TCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGCGAG AAGTGGAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAGTTCGGCAACAAGAC CATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCG GCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCG GCCCCAACACACCAACGGCACCATCACCCTGCCCTGCCGCATCAAGCAGATCATCAACCGCT GGGGCGCAAGGCCATGTACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCAACATC ACCGGCCTGCTGACCCGCGACGCGGCAAGGAGATCAGCAACACCACCGAGATCTTCCG CCCCGCGCGCGCACATGCGCGACAACTGGCGCAGCGAGCTGTACAAGTACAAGGTGGTGA AGATCGAGCCCTGGGCGTGGCCCCCACCAAGGCCAAGCGCCGCGTGGTGCAGCGCGAGAAG CGCGCCGTGACCCTGGGCGCATGTTCCTGGGCTTCCTGGGCGCCGCCGGCAGCACCATGGGC GCCCGCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGAGCGGCATCGTGCAGCAGCA GAACAACCTGCTGCGCGCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGGCA TCAAGCAGCTGCAGGCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTG GGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCCAGCTG ATCGACAACTACACCAACCTGATCTACACCCTGATCGAGGAGAGCCAGAACCAGCAGGAGAA GAACGAGCAGGAGCTGCAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTCGACATCA GCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTGGTGGGCCTGCGCA TCGTGTTCACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTCC AGACCCGCTTCCCCGCCCCCCGCGCCCCGACCGCCCCGAGGGCATCGAGGAGGGGGGGCGC GAGCGCGACCGCAGCAGCAGCCCCTGGTGCACGGCCTGCTGGCCCTGATCTGGGACGA CCTGCGCAGCCTGTGCCTGTTCAGCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCCG CATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGCTGC AGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCATCGCC GTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCCCAGCGCATCGGCCGCGCCTTCCTGCA CATCCCCGCCGCATCCGCCAGGCTTCGAGCGCCCCTGCTGTAACTCGAG

SEQ ID NO:10 ARG426-GLY431

GAATTCGCCACCATGGATGCAATGAAGAGGGGCTCTGCTGTGTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGTGAAGCTGACCCCCTGTGCGTG ACCCTGCACTGCACCTGAAGAACGCCACCAACACCAAGAGCAGCAACTGGAAGGAGAT GGACCGCGGCGAGATCAAGAACTGCAGCTTCAAGGTGACCACCAGCATCCGCAACAAGATGC AGAAGGAGTACGCCCTGTTCTACAAGCTGGACGTGGTGCCCATCGACAACGACAACACCAGC TACAAGCTGATCAACTGCAACACCAGCGTGATCACCCAGGCCTGCCCCAAGGTGAGCTTCGA GCCCATCCCATCACTACTGCGCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAA GTTCAACGGCAGCGCCCTGCACCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCC CCGTGGTGAGCACCCAGCTGCTGAACGGCAGCCTGGCCGAGGAGGGCGTGGTGATCCGC AGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGAGAT CAACTGCACCGCCCAACACACACCCGCAAGAGCATCACCATCGGCCCGGCCGCCCC TCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGCGAG AAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAGTTCGGCAACAAGAC CATCGTGTTCAAGCAGCAGCGGCGGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCG GCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCG GCCCCAACACCCACGGCACCATCACCCTGCCCTGCCGCATCAAGCAGATCATCAACCGC GGCGGCGGCAAGGCCATGTACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCAACAT CACCGGCCTGCTGACCCGCGACGGCGCAAGGAGATCAGCAACACCACCGAGATCTTCC GCCCCGGCGCGCGACATGCGCGACAACTGGCGCAGCGAGCTGTACAAGTACAAGTGGTG AAGATCGAGCCCTGGGCGTGGCCCCACCAAGGCCAAGCGCCGCGTGGTGCAGCGCGAGAA GCGCGCGTGACCCTGGGCGCCATGTTCCTGGGCTTCCTGGGCGCCGCCGGCAGCACCATGGG CGCCCGCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGAGCGGCATCGTGCAGCAGC AGAACAACCTGCTGCGCGCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGGC ATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCT GGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCCAGCT GATCGACAACTACACCAACCTGATCTACACCCTGATCGAGGAGGCCAGAACCAGCAGGAGA AGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTCGACATC AGCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTGGTGGGCCTGCGC ATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTC CAGACCCGCTTCCCCGCCCCCCGGGCCCCGACCGCCCGAGGGCATCGAGGAGGAGGGCGG CGAGCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTGGCCCTGATCTGGGACG ACCTGCGCAGCCTGTGCCTGTTCAGCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCC GCATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGCTG CAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCATCGC CGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCCCAGCGCATCGGCCGCGCCTTCCTGC ACATCCCCGCCGCATCCGCCAGGCTTCGAGCGCCCCTGCTGTAACTCGAG

SEQ ID NO:11 ARG426-GLY431B

GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGCCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTG ACCCTGCACTGCACCAACCTGAAGAACGCCACCAACACCAAGAGCAGCAACTGGAAGGAGAT GGACCGCGGGGAGATCAAGAACTGCAGCTTCAAGGTGACCACCAGCATCCGCAACAAGATGC AGAAGGAGTACGCCCTGTTCTACAAGCTGGACGTGGTGCCCATCGACAACGACAACACCAGC TACAAGCTGATCAACTGCAACACCAGCGTGATCACCCAGGCCTGCCCCAAGGTGAGCTTCGA GCCCATCCCATCCACTACTGCGCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAA GTTCAACGCCAGCGCCCTGCACCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCC CCGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCCGAGGAGGGCGTGGTGATCCGC AGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGAGAT CAACTGCACCCGCCCAACAACAACACCCGCAAGAGCATCACCATCGGCCCCGGCCGCCCT TCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGCGAG AAGTGGAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAGTTCGGCAACAAGAC CATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCG GCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCG GCCCCAACACCCAACGGCACCATCACCCTGCCCTGCCGCATCAAGCAGATCATCAACCGC GGCAGCGGCAAGGCCATGTACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCAACAT CACCGGCCTGCTGACCCGCGACGGCGCAAGGAGATCAGCAACACCACCGAGATCTTCC GCCCGGCGGCGACATGCGCGACAACTGGCGCAGCGAGCTGTACAAGTACAAGGTGGTG AAGATCGAGCCCTGGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGGTGCAGCGCGAGAA GCGCGCGTGACCTGGGCGCATGTTCCTGGGCTTCCTGGGCGCCGCCGGCAGCACCATGGG CGCCCGCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGAGCGGCATCGTGCAGCAGC AGAACAACCTGCTGCGCGCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGGC ATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCT GGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCCAGCT GATCGACAACTACACCAACCTGATCTACACCTGATCGAGGAGAGCCAGAACCAGCAGGAGA AGAACGAGCAGGAGCTGCAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTCGACATC AGCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCCTGGTGGGCCTGCGC ATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTC CAGACCCGCTTCCCCGCCCCCGCGGCCCCGACCGCCCGAGGGCATCGAGGAGGAGGGCGG CGAGCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTGGCCCTGATCTGGGACG ACCTGCGCAGCCTGTGCCTGTTCAGCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCC GCATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGCTG CAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCATCGC CGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCCCAGCGCATCGGCCGCGCCTTCCTGC ACATCCCCGCCGCATCCGCCAGGCTTCGAGCGCCCCTGCTGTAACTCGAG

SEQ ID NO:12 ARG426-LYS432

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTG ACCCTGCACTGCACCAACCTGAAGAACGCCACCAACACCAAGAGCAGCAACTGGAAGGAGAT GGACCGCGGCGAGATCAAGAACTGCAGCTTCAAGGTGACCACCAGCATCCGCAACAAGATGC AGAAGGAGTACGCCCTGTTCTACAAGCTGGACGTGGTGCCCATCGACAACGACAACACCAGC TACAAGCTGATCAACTGCAACACCAGCGTGATCACCCAGGCCTGCCCCAAGGTGAGCTTCGA GCCCATCCCATCCACTACTGCGCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAA GTTCAACGGCAGCGGCCCCTGCACCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCC CCGTGGTGAGCACCCAGCTGCTGAACGGCAGCCTGGCCGAGGAGGGCGTGGTGATCCGC AGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGAGAT CAACTGCACCGCCCAACAACACCCGCAAGAGCATCACCATCGGCCCGGCCGCCCCT TCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGCGAG AAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAGTTCGGCAACAAGAC CATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCG GCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCG GGCGGCAACAAGGCCATGTACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCAACAT CACCGGCTGCTGCTGACCCGCGACGGCGCAAGGAGATCAGCAACACCACCGAGATCTTCC GCCCGGCGGCGACATGCGCGACACTGGCGCGAGCTGTACAAGTACAAGTGGTG AAGATCGAGCCCTGGGCGTGGCCCCACCAAGGCCAAGCGCCGCGTGGTGCAGCGCGAGAA GCGCGCGTGACCCTGGGCGCATGTTCCTGGGCTTCCTGGGCGCCGCCGGCAGCACCATGGG CGCCGCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGAGCGGCATCGTGCAGCAGC AGAACAACCTGCTGCGCGCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGGC ATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCT GGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCCAGCT GATCGACAACTACACCAACCTGATCTACACCCTGATCGAGGAGAGCCAGAACCAGCAGGAGA AGAACGAGCAGGAGCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTCGACATC AGCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTGGTGGGCCTGCGC ATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTC CAGACCCGCTTCCCCGCCCCCGCGGCCCCGACCGCCCGAGGGCATCGAGGAGGAGGGCGG CGAGCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTGGCCCTGATCTGGGACG ACCTGCGCAGCCTGTGCCTGTTCAGCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCC GCATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGCTG CAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCATCGC CGTGGCCGAGGGCACCGCATCATCGAGGTGGCCCAGCGCATCGGCCGCCTTCCTGC ACATCCCCGCCGCATCCGCCAGGGCTTCGAGCGCCCCTGCTGTAACTCGAG

SEO ID NO:13 ASN425-LYS432

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTG ACCCTGCACTGCACCAACCTGAAGAACGCCACCAACACCAAGAGCAGCAACTGGAAGGAGAT GGACCGCGGCGAGATCAAGAACTGCAGCTTCAAGGTGACCACCAGCATCCGCAACAAGATGC AGAAGGAGTACGCCCTGTTCTACAAGCTGGACGTGGTGCCCATCGACAACGACAACACCAGC TACAAGCTGATCAACTGCAACACCAGCGTGATCACCCAGGCCTGCCCCAAGGTGAGCTTCGA GCCCATCCCATCACTACTGCGCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAA GTTCAACGGCAGCGCCCCTGCACCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCC CCGTGGTGAGCACCCAGCTGCTGAACGGCAGCCTGGCCGAGGAGGGCGTGGTGATCCGC AGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGAGAT CAACTGCACCGCCCAACAACAACACCCGCAAGAGCATCACCATCGGCCCCGGCCGCCCCT TCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGCGAG AAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAGTTCGGCAACAAGAC CATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCG GCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCG GCCCCAACACCCAACGGCACCATCACCCTGCCCTGCCGCATCAAGCAGATCATCAACGCCC CCAAGGCCATGTACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCAACATCACCGGCC TGCTGCTGACCCGCGACGGCGCAAGGAGATCAGCAACACCACCGAGATCTTCCGCCCCGGC GGCGGCGACATGCGCGACAACTGGCGCAGCGAGCTGTACAAGTACAAGGTGGTGAAGATCGA GCCCTGGGCGTGGCCCCACCAAGGCCAAGCGCCGCGTGGTGCAGCGCGAGAAGCGCGCCG TGACCCTGGGCGCATGTTCCTGGGCTTCCTGGGCGCCGCCGCAGCACCATGGGCGCCCGCA GCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGAGCGGCATCGTGCAGCAGCAGAACAAC CTGCTGCGCGCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGGCATCAAGCA GCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTGGGCATCT GGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCCAGCTGGAGCAAC CTACACCAACCTGATCTACACCTGATCGAGGAGAGCCAGAACCAGCAGGAGAAGAACGAGC AGGAGCTGCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTCGACATCAGCAAGTGG ACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTCCAGACCCGC TTCCCGCCCCCGGGCCCCGACGCCCCGAGGGCATCGAGGAGGAGGCGCGAGCGCGA CCGCGACCGCAGCACCCCTGGTGCACGCCTGCTGGCCCTGATCTGGGACGACCTGCGCAG GCTGCTGGGCCGCCGCGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGCTGCAGTACTGGA TCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCATCGCCGTGGCCGAG GGCACCGACCGCATCATCGAGGTGGCCCAGCGCATCGGCCGCCTTCCTGCACATCCCCCGC CGCATCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACTCGAG

SEQ ID NO:14 ILE424-ALA433

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTG ACCCTGCACTGCACCAACCTGAAGAACGCCACCAACACCAAGAGCAGCAACTGGAAGGAGAT GGACCGCGGCGAGATCAAGAACTGCAGCTTCAAGGTGACCACCAGCATCCGCAACAAGATGC AGAAGGAGTACGCCCTGTTCTACAAGCTGGACGTGGTGCCCATCGACAACGACAACACCAGC TACAAGCTGATCAACTGCAACACCAGCGTGATCACCCAGGCCTGCCCCAAGGTGAGCTTCGA GCCATCCCATCACTACTGCGCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAA GTTCAACGGCAGCGCCCCTGCACCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCC CCGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCCGAGGAGGGCGTGGTGATCCGC AGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGAGAT CAACTGCACCGCCCAACAACACCCGCAAGAGCATCACCATCGGCCCCGGCCGCCCCT TCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGCGAG AAGTGGAACACACCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAGTTCGGCAACAAGAC ${\tt CATCGTGTTCAAGCAGAGCAGCGGGGGGGGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCG}$ GCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCG GCCCAACAACACCAACGCACCATCACCTGCCTGCCGCATCAAGCAGATCATCGGCGGC GCCATGTACGCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCAACATCACCGGCCTGCTG CTGACCCGCGACGCGCAAGGAGATCAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGG CGACATGCGCGACAACTGGCGCAGCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCC TGGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGGTGCAGCGCGAGAAGCGCGCGTGACC CTGGGCGCCATGTTCCTGGGCTTCCTGGGCGCCGCGGCACCATGGGCGCCCGCAGCCTG ACCCTGACCGTGCAGGCCCGCCAGCTGCTGAGCGGCATCGTGCAGCAGCAGCAGCAACCACCTGCT GCGCGCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGC AGGCCGCGTGCTGGCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGC TGCAGCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAG CCTGGACCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGGCGCGAGATCGACAACTACA CCAACCTGATCTACACCCTGATCGAGGAGAGCCAGAACCAGCAGGAGAACGAGCAGGA GCTGCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTCGACATCAGCAAGTGGCTGT GGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTGGTGGGCCTGCGCATCGTGTTCACCG TGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTCCAGACCCGCTTCC CCGCCCCGCGGCCCGACGCCCCGAGGGCATCGAGGAGGAGGGCGCGAGCGCGACCGC GACCGCAGCAGCCCCTGGTGCACGGCCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCCTG CTGGGCCGCGCGGGTGGGAGGCCTGAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCA GGAGCTGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCATCGCCGTGGCCGAGGGCA CCGACCGCATCATCGAGGTGGCCCAGCGCATCGGCCGCCTTCCTGCACATCCCCCGCCGCA TCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACTCGAG

SEQ ID NO:15 ILE423-MET434

GAATTCGCCACCATGGATGCAATGAAGAGGGCTCTGCTGTTGTGCTGCTGTTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGCCACCACCCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTG ACCTGCACTGCACCAACCTGAAGAACGCCACCAACACCAAGAGCAGCAACTGGAAGGAGAT GGACCGCGGCGAGATCAAGAACTGCAGCTTCAAGGTGACCACCAGCATCCGCAACAAGATGC AGAAGGAGTACGCCTGTTCTACAAGCTGGACGTGGTGCCCATCGACAACGACAACACCAGC TACAAGCTGATCAACTGCAACACCAGCGTGATCACCCAGGCCTGCCCCAAGGTGAGCTTCGA GCCCATCCCATCCACTACTGCGCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAA GTTCAACGGCAGCGCCCCTGCACCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCC CCGTGGTGAGCACCCAGCTGCTGAACGGCAGCCTGGCCGAGGAGGGCGTGGTGATCCGC AGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGAGAT CAACTGCACCGCCCAACACACACCCGCAAGAGCATCACCATCGGCCCGGCCGCCCCC TCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGCGAG AAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAGTTCGGCAACAAGAC CATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCG GCGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCG GCCCCAACACCCAACGGCACCATCACCCTGCCCTGCCGCATCAAGCAGATCGGCGGCATG TACGCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCAACATCACCGGCCTGCTGACC CGCGACGGCGGCAAGGAGATCAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGGCGACAT GCGCGACAACTGGCGCAGCGAGCTGTACAAGTACAAGTGGTGAAGATCGAGCCCCTGGGCG TGGCCCCACCAAGGCCAAGCGCCGCGTGGTGCAGCGCGAGAAGCGCGCCGTGACCCTGGGC GCCATGTTCCTGGGCTTCCTGGGCGCCGCCGCAGCACCATGGGCGCCCGCAGCCTGACCCTG ACCGTGCAGCCCGCCAGCTGCTGAGCGGCATCGTGCAGCAGCAGCAGCACCAGCTGCTGCGCGC CATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGGCCATCAAGCAGCTGCAGGCCC GCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCAGC GGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGA CCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGCGAGATCGACAACTACACCAACC TGATCTACACCCTGATCGAGGAGAGCCAGAACCAGCAGGAGAACAACGAGCAGGAGCTGCTG GAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTCGACATCAGCAAGTGGCTGTGGTACAT CAAGATCTTCATCATGATCGTGGGCGGCCTGGTGGGCCTGCGCATCGTGTTCACCGTGCTGAG CATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCC AGCAGCCCCTGGTGCACGGCCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTG CGCCGCGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCT GAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCATCGCCGTGGCCGAGGGCACCGACC GCATCATCGAGGTGGCCCAGCGCATCGGCCGCCTTCCTGCACATCCCCGCCGCATCCGCC AGGGCTTCGAGCGCCCCTGCTGTAACTCGAG

SEQ ID NO:16 GLN422-TYR435

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGCCACCACCCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTG ACCTGCACTGCACCAACCTGAAGAACGCCACCAACACCAAGAGCAGCAACTGGAAGGAGAT GGACCGCGGCGAGATCAAGAACTGCAGCTTCAAGGTGACCACCAGCATCCGCAACAAGATGC AGAAGGAGTACGCCCTGTTCTACAAGCTGGACGTGGTGCCCATCGACAACGACAACACCAGC TACAAGCTGATCAACTGCAACACCAGCGTGATCACCCAGGCCTGCCCCAAGGTGAGCTTCGA GCCCATCCCATCACTACTGCGCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAA GTTCAACGCCAGCGCCCTGCACCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCC CCGTGGTGAGCACCCAGCTGCTGCTGAACGCCAGCCTGGCCGAGGAGGGCGTGGTGATCCGC AGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGAGAT CAACTGCACCGCCCAACACACACCCGCAAGAGCATCACCATCGGCCCGGCCGCCCC TCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGCGAG AAGTGGAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAGTTCGGCAACAAGAC CATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCG GCGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCG GCCCCAACACACCAACGCACCATCACCCTGCCCTGCCGCATCAAGCAGGGCGGCTACGCC CCCCCATCCGCGGCCAGATCCGCTGCAGCAGCAACATCACCGGCCTGCTGCTGACCCGCGAC GGCGGCAAGGAGATCAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGGCGACATGCGCGA CAACTGGCGCAGCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCC CCACCAAGGCCAAGCGCGCGTGGTGCAGCGCGAGAAGCGCGCGTGACCCTGGGCGCCATG TTCCTGGGCTTCCTGGGCGCCGCCGCAGCACCATGGGCGCCCGCAGCCTGACCCTG CAGGCCGCCAGCTGCTGAGCGGCATCGTGCAGCAGCAGAACAACCTGCTGCGCGCCATCGA GGCCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGC TGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCAGCGGCAAG CTGATCTGCACCACCGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAGAT CTGGAACAACATGACCTGGATGGAGTGGGAGCGCGAGATCGACAACTACACCAACCTGATCT ACACCCTGATCGAGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGAGCT GGACAGTGGGCCAGCCTGTGGAACTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGA TCTTCATCATGATCGTGGGCGGCCTGGTGGGCCTGCGCATCGTGTTCACCGTGCTGAGCATCG TGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCCGCG CCCCTGGTGCACGGCCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCAG CGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGA ATCGAGGTGGCCCAGCGCATCGGCCGCCCTTCCTGCACATCCCCCGCCGCATCCGCCAGGGC TTCGAGCGCCCCTGCTGTAACTCGAG

SEO ID NO:17 GLN422-TYR435B

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCTGCGTGAAGCTGACCCCCCTGTGCGTG ACCCTGCACTGCACCAACCTGAAGAACGCCACCAACACCAAGAGCAGCAACTGGAAGGAGAT GGACCGCGGCGAGATCAAGAACTGCAGCTTCAAGGTGACCACCAGCATCCGCAACAAGATGC AGAAGGAGTACGCCCTGTTCTACAAGCTGGACGTGGTGCCCATCGACAACGACAACACCAGC TACAAGCTGATCAACTGCAACACCAGCGTGATCACCCAGGCCTGCCCCAAGGTGAGCTTCGA GCCATCCCATCACTACTGCGCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAA GTTCAACGGCAGCGGCCCTGCACCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCC CCGTGGTGAGCACCCAGCTGCTGAACGGCAGCCTGGCCGAGGAGGGCGTGGTGATCCGC AGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGAGAT CAACTGCACCGCCCCAACACACACCCCGCAAGAGCATCACCATCGGCCCCGGCCGCCCC TCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGCGAG AAGTGGAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAGTTCGGCAACAAGAC CATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCG GCGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCG GCCCAACAACACCAACGGCACCATCACCCTGCCTGCCGCATCAAGCAGGCCCCCTACGCCC CCCCATCGCGGCCAGATCCGCTGCAGCAGCAACATCACCGGCCTGCTGCTGACCCGCGACG GCGGCAAGGAGATCACCACCGAGATCTTCCGCCCCGGCGCGCGACATGCGCGAC AACTGGCGCAGCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCC CACCAAGGCCAAGCGCGCGTGGTGCAGCGCGAGAAGCGCGCGTGACCCTGGGCGCCATGT TCCTGGGCTTCCTGGGCGCCGCGGCAGCACCATGGGCGCCCGCAGCCTGACCCTGACCGTGC AGGCCGCCAGCTGCTGAGCGGCATCGTGCAGCAGCAGCAACCTGCTGCGCGCCCATCGAG GCCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGCT GGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCAGCGGCAAGC TGATCTGCACCACCGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAGATC TGGAACAACATGACCTGGATGGAGTGGGAGCGCGAGATCGACAACTACACCAACCTGATCTA CACCTGATCGAGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTG GACAAGTGGGCCAGCTGTGGAACTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGAT CTTCATCATGATCGTGGCGGCCTGGTGGGCCTGCGCATCGTGTTCACCGTGCTGAGCATCGT GAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCCGCGG CCCTGGTGCACGCCTGCTGGCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCAGC TACCACCGCTGCGCGACCTGATCCTGATCGCCGCCGCATCGTGGAGCTGCTGGGCCGCCGC GGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAA CGAGGTGGCCCAGCGCATCGGCCGCGCCTTCCTGCACATCCCCCGCCGCATCCGCCAGGGCTT CGAGCGCCCTGCTGTAACTCGAG

57 / 65

SEQ ID NO:18: LEU122-SER199; ARG426-GLY431

GAATTCGCCACCATGGATGCAATGAAGAGGGGCTCTGCTGTTGTGCTGCTGTTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGGGCAACAGCGTGAT CACCCAGGCCTGCCCCAAGGTGAGCTTCGAGCCCATCCCCATCACTGCGCCCCGCCGG CTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCCTGCACCAACGTGA GCACCGTGCAGTGCACCCACGGCATCCGCCCCGTGGTGAGCACCCAGCTGCTGCTGAACGGC AGCCTGGCCGAGGAGGGCGTGGTGATCCGCAGCGAGAACTTCACCGACAACGCCAAGACCAT CATCGTGCAGCTGAAGGAGAGCGTGGAGATCAACTGCACCGCCCCAACAACACCCCGCA AGAGCATCACCATCGGCCCGGCCGCCCTTCTACGCCACCGGCGACATCATCGGCGACATCC GCCAGGCCCACTGCAACATCAGCGGCGAGAAGTGGAACAACACCCTGAAGCAGATCGTGACC AAGCTGCAGGCCCAGTTCGGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGGCGGCGACCC CGAGATCGTGATGCACAGCTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACCCAGCT GTTCAACAGCACCTGGAACAACACCATCGGCCCCAACAACACCAACGGCACCATCACCCTGC CCTGCCGCATCAAGCAGATCATCAACCGCGGCGGCGAAGGCCATGTACGCCCCCCCATCC GCGGCCAGATCCGCTGCAGCAGCAACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAAG GAGATCAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGGCGACATGCGCGACAACTGGCG CAGCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCACCAAGG CCAAGCGCCGCGTGCCGCGAGAAGCGCGCGTGACCCTGGGCGCATGTTCCTGGGC TTCCTGGGCGCCGCCGCAGCACCATGGGCGCCCGCAGCCTGACCCTGACCGTGCAGGCCCGC CAGCTGCTGAGCGGCATCGTGCAGCAGCAGCACCACCTGCTGCGCGCCCATCGAGGCCCAGCA GCACCTGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGG AGCGCTACCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGC ACCACCGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAGATCTGGAACAA CATGACCTGGATGGAGTGGGAGCGCGAGATCGACAACTACACCAACCTGATCTACACCCTGA TCGAGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTG GGCCAGCCTGTGGAACTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCAT GATCGTGGGCGCCTGGTGGGCCTGCGCATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGT GCGCCAGGGCTACAGCCCCTGAGCTTCCAGACCGCTTCCCCGCCCCCCGCGGCCCGACCG CCCCGAGGCATCGAGGAGGGGGGCGCGAGCGCGACCGCAGCAGCCCCTGGTGC ACGGCCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCAGCTACCACCGCC TGCGCGACCTGATCCTGATCGCCGCCCGCATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGG CCCTGAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTG AGCCTGTTCGACGCCATCGCCATCGCCGTGGCCGAGGGCACCGCATCATCGAGGTGGCC CAGCGCATCGGCCGCCTTCCTGCACATCCCCCGCCGCATCCGCCAGGGCTTCGAGCGCGCC **CTGCTGTAACTCGAG**

SEQ ID NO:19 LEU122-SER199; ARG426-LYS432

GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGCCACCACCCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGGGCAACAGCGTGAT CACCCAGGCCTGCCCCAAGGTGAGCTTCGAGCCCATCCCATCCACTACTGCGCCCCGCCGG CTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCCTGCACCAACGTGA GCACCGTGCAGTGCACCCACGGCATCCGCCCCGTGGTGAGCACCCAGCTGCTGCAACGGC AGCCTGGCCGAGGGGGCGTGGTGATCCGCAGCGAGAACTTCACCGACAACGCCAAGACCAT CATCGTGCAGCTGAAGGAGAGCGTGGAGATCAACTGCACCCGCCCAACAACAACACCCCGCA AGAGCATCACCATCGGCCCGGCCGCCCTTCTACGCCACCGGCGACATCATCGGCGACATCC GCCAGGCCCACTGCAACATCAGCGGCGAGAAGTGGAACAACACCCTGAAGCAGATCGTGACC AAGCTGCAGGCCCAGTTCGGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGGCGGCGACCC CGAGATCGTGATGCACAGCTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACCCAGCT GTTCAACAGCACCTGGAACACCACCGGCCCCAACACCACCGGCACCATCACCCTGC CCTGCCGCATCAAGCAGATCATCAACCGCGGCGCAACAAGGCCATGTACGCCCCCCCATCC GCGCCAGATCCGCTGCAGCAGCAACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAAG GAGATCAGCAACACCACGAGATCTTCCGCCCCGGCGGCGGCGACATGCGCGACAACTGGCG CAGCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCACCAAGG CCAAGCGCCGCGTGGTGCAGCGCGAGAAGCGCGCGTGACCCTGGGCGCCATGTTCCTGGGC TTCCTGGGCGCCGCGGCAGCACCATGGGCGCCCGCAGCCTGACCCTGACCGTGCAGGCCCGC CAGCTGCTGAGCGCATCGTGCAGCAGCAGCACCACCTGCTGCGCGCCATCGAGGCCCAGCA GCACCTGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGG AGCGCTACCTGAAGGACCAGCAGCTGCTGGGCCATCTGGGGCTGCAGCGGCAAGCTGATCTGC ACCACCGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAGATCTGGAACAA CATGACCTGGATGGAGTGGGAGCCCGAGATCGACAACTACACCAACCTGATCTACACCCTGA TCGAGGAGACCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTG GGCCAGCCTGTGGAACTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCAT GATCGTGGGCGGCCTGGGGCCTGCGCATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGT GCGCCAGGGCTACAGCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCCGCGGCCCCGACCG ACGCCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCAGCTACCACCGCC TGCGCGACCTGATCCTGATCGCCGCCCGCATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGG CCCTGAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTG CAGCGCATCGGCCGCCTTCCTGCACATCCCCGCCGCATCCGCCAGGGCTTCGAGCGCGCC **CTGCTGTAACTCGAG**

SEO ID NO: 20: LEU122-SER199; TRP427-GLY431

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGGGCAACAGCGTGAT CACCCAGGCCTGCCCCAAGGTGAGCTTCGAGCCCATCCCCATCCACTACTGCGCCCCGCCGG CTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCCTGCACCAACGTGA GCACCGTGCAGTGCACCCACGGCATCCGCCCCGTGGTGAGCACCCAGCTGCTGCAACGGC AGCCTGGCCGAGGGGGCGTGGTGATCCGCAGCGAGACTTCACCGACAACGCCAAGACCAT CATCGTGCAGCTGAAGGAGAGCGTGGAGATCAACTGCACCCGCCCAACAACACCCCGCA AGAGCATCACCATCGGCCCGGCCGCCCTTCTACGCCACCGGCGACATCATCGGCGACATCC GCCAGGCCCACTGCAACATCAGCGGCGAGAAGTGGAACAACACCCTGAAGCAGATCGTGACC AAGCTGCAGGCCCAGTTCGGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGGCGGCGACCC CGAGATCGTGATGCACAGCTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACCCAGCT GTTCAACAGCACCTGGAACAACACCATCGGCCCCAACAACACCAACGGCACCATCACCCTGC CCTGCCGCATCAAGCAGATCATCAACCGCTGGGGCGGCAAGGCCATGTACGCCCCCCCATCC GCGCCAGATCCGCTGCAGCAGCAACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAAG GAGATCAGCAACACCACGAGATCTTCCGCCCGGCGGCGACATGCGCGACAACTGGCG CAGCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCACCAAGG CCAAGCGCCGCGTGGTGCAGCGCGAGAAGCGCGCCGTGACCTGGGCGCCATGTTCCTGGGC TTCTGGGCGCCGCAGCACCATGGGCGCCCGCAGCCTGACCCTGACCGTGCAGGCCGC CAGCTGCTGAGCGGCATCGTGCAGCAGCAGCACCACCTGCTGCGCGCCCATCGAGGCCCAGCA GCACCTGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGG AGCGCTACCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGC ACCACCGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAGATCTGGAACAA CATGACCTGGATGGAGTGGGAGCGCGAGATCGACAACTACACCAACCTGATCTACACCCTGA TCGAGGAGACCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTG GGCCAGCCTGTGGAACTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCAT GATCGTGGGCGGCTGGGGCCTGCGCATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGT GCGCCAGGGCTACAGCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCCGCGGCCCCGACCG CCCCGAGGGCATCGAGGAGGGCGGCGAGCGCGACCGCAGCAGCCCCCTGGTGC ACGGCCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCAGCTACCACCGCC TGCGCGACCTGATCCTGATCGCCGCCCGCATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGG CCCTGAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTG CAGCGCATCGGCCGCCTTCCTGCACATCCCCCGCCGCATCCGCCAGGGCTTCGAGCGCGCC CTGCTGTAACTCGAG

60 / 65

SEQ ID NO:21 LYS121-VAL200; ASN425-LYS432

GAATTCGCCACCATGGATGCAATGAAGAGGGGCTCTGCTGTGTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGGCCCCCGTGATCACCCA GGCCTGCCCAAGGTGAGCTTCGAGCCCATCCCCATCCACTACTGCGCCCCGGCCTGGCTTCGC CATCCTGAAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCCTGCACCAACGTGAGCACCG TGCAGTGCACCCACGGCATCCGCCCCGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGG CCGAGGAGGCGTGGTGATCCGCAGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTG CAGCTGAAGGAGAGCGTGGAGATCAACTGCACCGCCCCAACAACAACACCCGCAAGAGCAT CACCATCGGCCCGGCCGCCTTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGC CCACTGCAACATCAGCGGCGAGAAGTGGAACACACCCTGAAGCAGATCGTGACCAAGCTGC AGGCCCAGTTCGGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATC GTGATGCACAGCTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAAC AGCACCTGGAACACACCATCGGCCCCAACAACACCAACGGCACCATCACCCTGCCCTGCCG CATCAAGCAGATCATCAACGCCCCCAAGGCCATGTACGCCCCCCCATCCGCGGCCAGATCCG CTGCAGCAGCAACATCACCGGCCTGCTGCTGACCCGCGACGGCGAAGGAGATCAGCAACA CCACCGAGATCTTCCGCCCGGCGGCGGCGACATGCGCGACAACTGGCGCAGCGAGCTGTAC AAGTACAAGGTGGTGAAGATCGAGCCCTGGGCGTGGCCCCACCAAGGCCAAGCGCCGCGT GGTGCAGCGCGAGAAGCGCGCCGTGACCCTGGGCGCCATGTTCCTGGGCTTCCTGGGCGCCGC CGGCAGCACCATGGGCGCCCGCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGAGCG GCATCGTGCAGCAGCAGCACCACCTGCTGCGCGCCCATCGAGGCCCAGCAGCACCTGCTGCAG CTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGAA GGACCAGCAGCTGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGTGC CCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAGATCTGGAACAACATGACCTGGATG GAGTGGGAGCGCGAGATCGACAACTACACCAACCTGATCTACACCCTGATCGAGGAGAGCCA GAACCAGCAGGAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCAGCCTGTGG AACTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGC CTGGTGGGCCTGCGCATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTAC AGCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCGCGGCCCCGACCGCCCCGAGGGCATC GAGGAGGAGGCGCGACCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTGGC CCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCAGCTACCACCGCCTGCGCGACCTGAT CCTGATCGCCGCCCGCATCGTGGAGCTGCTGGGCCGCCGCGGCTGGAGGCCCTGAAGTACTG GGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTCGACG CCATCGCCATCGCCGTGGCCGAGGGCACCGACCGCATCGAGGTGGCCCAGCGCATCGGC CGCGCCTTCCTGCACATCCCCGCCGCATCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACTC GAG

61 / 65

SEO ID NO:22 VAL120-ILE201; ILE 424-ALA433

GAATTCGCCACCATGGATGCAATGAAGAGGGGCTCTGCTGTGTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGGGCGGCATCACCCAGGCCTG CCCCAAGGTGAGCTTCGAGCCCATCCCCATCCACTACTGCGCCCCGCCGGCTTCGCCATCCT GAAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCCTGCACCAACGTGAGCACCGTGCAGT GCACCCACGCATCCGCCCCGTGGTGAGCACCCAGCTGCTGCAACGGCAGCCTGGCCGAG GAGGGCGTGGTGATCCGCAGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCT GAAGGAGAGCGTGGAGATCAACTGCACCCGCCCAACAACAACACCCCGCAAGAGCATCACCA TCGGCCCGGCCGCCTTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACT GCAACATCAGCGGCGAGAAGTGGAACACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCC CAGTTCGGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGAT GCACAGCTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCAC CTGGAACACACCATCGGCCCCAACACACCAACGGCACCATCACCCTGCCCTGCCGCATCA AGCAGATCATCGGCGCCCCATGTACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGC AACATCACCGGCCTGCTGACCCGCGACGGCGCAAGGAGATCAGCAACACCACCGAGAT CTTCCGCCCGGCGGCGCGACATGCGCGACAACTGGCGCAGCGAGCTGTACAAGTACAAGG TGGTGAAGATCGAGCCCTGGGCGTGGCCCCACCAAGGCCAAGCGCCGCGTGGTGCAGCGC GAGAAGCGCCCTGACCCTGGGCGCCATGTTCCTGGGCTTCCTGGGCGCCGCCGGCAGCACC ATGGGCGCCGCAGCCTGACCTGACCGTGCAGGCCCGCCAGCTGCTGAGCGGCATCGTGCA GCAGCAGAACAACCTGCTGCGCCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGT GGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAG CTGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCC CGAGATCGACAACTACACCAACCTGATCTACACCCTGATCGAGGAGAGCCAGAACCAGCAGG AGAAGAACGAGCAGGAGCTGCTGGACCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTCGAC ATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTGGTGGGCCTG CGCATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGC TTCCAGACCCGCTTCCCCGCCCCCCGGGCCCCGACCGCCCCGAGGGCATCGAGGAGGAGGG CGGCGAGCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTGGCCCTGATCTGGG ACGACCTGCGCAGCCTGTTCCAGCTACCACCGCCTGCGCGACCTGATCCTGATCGCCG CCCGCATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTG CTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCATC GCCGTGGCCGAGGGCACCGCATCATCGAGGTGGCCCAGCGCATCGGCCGCCTTCCT GCACATCCCCGCCGCATCCGCCAGGGCTTCGAGCGCCCTGCTGTAACTCGAG

62 / 65

SEQ ID NO:23: VAL120-ILE201B; ILE424-ALA433

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGCCACCACCCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGCCCGGCATCACCCAGGCCTGC CCCAAGGTGAGCTTCGAGCCCATCCCCATCCACTACTGCGCCCCGCCGGCTTCGCCATCCTG AAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCCTGCACCAACGTGAGCACCGTGCAGTG CACCCACGGCATCCGCCCGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCCGAGG AGGGCGTGGTGATCCGCAGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCTG AAGGAGAGCGTGGAGATCAACTGCACCCGCCCAACACAACACCCCGCAAGAGCATCACCAT CGGCCCGGCCGCCTTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTG CAACATCAGCGGCGAGAAGTGGAACACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCC AGTTCGGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGATG CACAGCTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACC TGGAACAACACCATCGGCCCCAACAACACCCAACGGCACCATCACCCTGCCCTGCCGCATCAA GCAGATCATCGGCGGCGCCATGTACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCA ACATCACCGGCCTGCTGACCCGCGACGGCGGCAAGGAGATCAGCAACACCACCGAGATC TTCCGCCCGGCGGCGACATGCGCGACACTGGCGCAGCGAGCTGTACAAGTACAAGGT GGTGAAGATCGAGCCCTGGGCGTGGCCCCACCAAGGCCAAGCGCCGCGTGGTGCAGCGCG AGAAGCGCGCGTGACCCTGGGCGCCATGTTCCTGGGCTTCCTGGGCGCCGCCGCAGCACCA TGGGCGCCGCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGAGCGGCATCGTGCAG CAGCAGAACAACCTGCTGCGCCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGTG GGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGC TGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCCA GAGATCGACAACTACACCAACCTGATCTACACCCTGATCGAGGAGAGCCAGAACCAGCAGGA GAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTCGACA GCATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCT TCCAGACCCGCTTCCCCGCCCCCGCGGCCCCGACCGCCCCGAGGGCATCGAGGAGGAGGGC GGCGAGCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTGGCCCTGATCTGGGA CGACCTGCGCAGCCTGTGCCTGTTCAGCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGC CCGCATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGC TGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCATC GCCGTGGCCGAGGGCACCGCATCATCGAGGTGGCCCAGCGCATCGGCCGCGCCTTCCT GCACATCCCCGCCGCATCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACTCGAG

SEQ ID NO:24 VAL120-THR202; ILE424-ALA433

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGGGCGCGCCACCCAGGCCTG CCCCAAGGTGAGCTTCGAGCCCATCCCCATCCACTACTGCGCCCCGCCGGCTTCGCCATCCT GAAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCCTGCACCAACGTGAGCACCGTGCAGT GCACCCACGGCATCCGCCCCGTGGTGAGCACCCAGCTGCTGAACGGCAGCCTGGCCGAG GAGGGCGTGGTGATCCGCAGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCT GAAGGAGAGCGTGGAGATCAACTGCACCCGCCCCAACAACACCCCGCAAGAGCATCACCA TCGGCCCCGGCCGCCTTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACT GCAACATCAGCGGCGAGAAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCC CAGTTCGGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGAT GCACAGCTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCAC CTGGAACAACACCATCGGCCCCAACAACACCAACGGCACCATCACCCTGCCCTGCCGCATCA AGCAGATCATCGGCGGCGCCATGTACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGC AACATCACCGGCCTGCTGACCCGCGACGGCGGCAAGGAGATCAGCAACACCACCGAGAT CTTCCGCCCGGCGGCGCGACATGCGCGACAACTGGCGCAGCGAGCTGTACAAGTACAAGG TGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCACCAAGGCCAAGCGCCGCGTGGTGCAGCGC GAGAAGCGCCCTGACCCTGGGCGCCATGTTCCTGGGCTTCCTGGGCGCCGCCGCCAGCACC ATGGGCGCCCGCAGCCTGACCCTGACCGTGCAGCCCCAGCTGCTGAGCGGCATCGTGCA GCAGCAGAACAACCTGCTGCGCGCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGT GGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAG CTGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCC CGAGATCGACAACTACACCAACCTGATCTACACCCTGATCGAGGAGAGCCAGAACCAGCAGG AGAAGAACGAGCAGGAGCTGCAGGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTCGAC ATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTGGTGGGCCTG CGCATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGC TTCCAGACCCGCTTCCCCGCCCCCCGCGGCCCCGACCGCCCCGAGGGCATCGAGGAGGAGGG CGGCGAGCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTGGCCCTGATCTGGG ACGACCTGCGCAGCCTGTGCCTGTTCAGCTACCACCGCCTGCGCGACCTGATCCTGATCGCCG CCCGCATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTG CTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCATC GCCGTGGCCGAGGGCACCGCATCATCGAGGTGGCCCAGCGCATCGGCCGCCCTTCCT GCACATCCCCGCCGCATCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACTCGAG

SEQ ID NO:25 VAL127-ASN195

GAATTCGCCACCATGGATGCAATGAAGAGGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGCCACCACCCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTG GGGGCAGGGAACTGCAACACCAGCGTGATCACCCAGGCCTGCCCCAAGGTGAGCTTCGAGCC CATCCCCATCCACTACTGCGCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAAGTT CAACGCAGCGCCCTGCACCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCCCG TGGTGAGCACCCAGCTGCTGAACGGCAGCCTGGCCGAGGAGGGCGTGGTGATCCGCAGC GAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGAGATCAA CTGCACCGCCCAACAACAACACCCGCAAGAGCATCACCATCGGCCCCGGCCGCCCTTCTA CGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGCGAGAAGT GGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAGTTCGGCAACAAGACCATC GTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCGGCGG CGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCGGCCC CAACAACACCAACGGCACCATCACCCTGCCCTGCCGCATCAAGCAGATCATCAACCGCTGGC AGGAGGTGGCAAGGCCATGTACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCAAC ATCACCGGCCTGCTGACCCGCGACGGCGCGAGGAGATCAGCAACACCACCGAGATCTT CCGCCCGGCGCGCGACATGCGCGACAACTGGCGCAGCGAGCTGTACAAGTACAAGTGG TGAAGATCGAGCCCTGGGCGTGGCCCCACCAAGGCCAAGCGCCGCGTGGTGCAGCGCGAG AAGCGCGCGTGACCCTGGGCGCCATGTTCCTGGGCTTCCTGGGCGCCGCCGGCAGCACCATG GGCGCCGCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGAGCGGCATCGTGCAGCA GCAGAACAACCTGCTGCGCGCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGTGGG GCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTG CTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCCAG AGATCGACAACTACACCAACCTGATCTACACCCTGATCGAGGAGGCCAGAACCAGCAGGAG AAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTCGACAT CATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTT CCAGACCCGCTTCCCCGCCCCCGGGCCCCGACGCCCCGAGGGCATCGAGGAGGAGGGCG GCGAGCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTGGCCCTGATCTGGGAC GACCTGCGCAGCCTGTGCCTGTTCAGCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCC CGCATCGTGGAGCTGCTGGGCCGCCGCGGCTGGAGGCCCTGAAGTACTGGGGCAACCTGCT GCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCATCG CCGTGGCCGAGGGCACCGCATCATCGAGGTGGCCCAGCGCATCGGCCGCGCTTCCTGC ACATCCCCGCCGCATCCGCCAGGGCTTCGAGCGCCCCTGCTGTAACTCGAG

65 / 65

SEQ ID NO:26 VAL127-ASN195; ARG426-GLY431

GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTG GGGGCAGGGAACTGCAACACCAGCGTGATCACCCAGGCCTGCCCCAAGGTGAGCTTCGAGCC CATCCCCATCCACTACTGCGCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAAGTT CAACGCAGCGCCCTGCACCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCCCG TGGTGAGCACCCAGCTGCTGAACGGCAGCCTGGCCGAGGAGGGCGTGGTGATCCGCAGC GAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGAGATCAA CTGCACCGCCCAACAACACCCGCAAGAGCATCACCATCGGCCCCGGCCGCCCTTCTA CGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGCGAGAAGT GGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAGTTCGGCAACAAGACCATC GTGTTCAAGCAGAGCAGCGGGGGGACCCCGAGATCGTGATGCACAGCTTCAACTGCGGCGG CGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCGGCCC CAACAACACCAACGCACCATCACCCTGCCCTGCCGCATCAAGCAGATCATCAACCGCGGCG GCGGCAAGGCCATGTACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCAACATCACC GGCCTGCTGACCCGCGACGGCGGCAAGGAGATCAGCAACACCACCGAGATCTTCCGCCC CGGGGCGCGACATGCGCGACAACTGGCGCAGCGAGCTGTACAAGTACAAGTGGTGAAG ATCGAGCCCTGGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGGTGCAGCGCGAGAAGCG CGCCGTGACCCTGGGCGCATGTTCCTGGGCTTCCTGGGCGCCGCCGGCAGCACCATGGGCGC CCGCAGCCTGACCCTGCAGGCCCGCCAGCTGCTGAGCGGCATCGTGCAGCAGCAGA ACAACCTGCTGCGCGCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGGCATC AAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTGGG CATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCCAGCTGGA GACAACTACACCAACCTGATCTACACCCTGATCGAGGAGAGCCAGAACCAGCAGGAGAAGAA CGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTCGACATCAGCA AGTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTGGTGGGCCTGCGCATCG TGTTCACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTCCAGA CGCGACCGCGACCGCAGCCCCCTGGTGCACGGCCTGCTGGCCCTGATCTGGGACGACCTG GTGGAGCTGCTGGGCCGCGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGCTGCAGTA CTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCATCGCCGTGG CCGAGGCCACCGCATCATCGAGGTGGCCCAGCGCATCGGCCGCCTTCCTGCACATCC CCCGCCGCATCCGCCAGGGCTTCGAGCGCCCCTGCTGTAACTCGAG

SEQUENCE LISTING

<110> Chiron Corporation
<120> MODIFIED HIV ENV POLYPEPTIDES
<130> 1605.100
<140>
<141>
<160> 26
<170> PatentIn Ver. 2.0

<210> 1

<211> 856

<212> PRT

<213> Human immunodeficiency virus

<400> 1

Met Arg Val Lys Glu Lys Tyr Gln His Leu Trp Arg Trp Gly Trp Arg 1 5 10 15

Trp Gly Thr Met Leu Leu Gly Met Leu Met Ile Cys Ser Ala Thr Glu 20 25 30

Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala
35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu
50 60

Val His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn 65 70 75 80

Pro Gln Glu Val Val Leu Val Asn Val Thr Glu Asn Phe Asn Met Trp 85 90 95

Lys Asn Asp Met Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp
100 105 110

Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Ser 115 120 125

Gly Arg Met Ile Met Glu Lys Gly Glu Ile Lys Asn Cys Ser Phe Asn 145 150 155 160

Ile Ser Thr Ser Ile Arg Gly Lys Val Gln Lys Glu Tyr Ala Phe Phe 165 170 175

Tyr Lys Leu Asp Ile Ile Pro Ile Asp Asn Asp Thr Thr Ser Tyr Lys 180 185 190

1

Leu Thr Ser Cys Asn Thr Ser Val Ile Thr Gln Ala Cys Pro Lys Val

- Ser Phe Glu Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Phe Ala 210 215 220
- Ile Leu Lys Cys Asn Asn Lys Thr Phe Asn Gly Thr Gly Pro Cys Thr 225 230 235 240
- Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Arg Pro Val Val Ser
- Thr Gln Leu Leu Asn Gly Ser Leu Ala Glu Glu Glu Val Val Ile
 260 265 270
- Arg Ser Val Asn Phe Thr Asp Asn Ala Lys Thr Ile Ile Val Gln Leu 275 280 285
- Asn Thr Ser Val Glu Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg 290 295 300
- Lys Arg Ile Arg Ile Gln Arg Gly Pro Gly Arg Ala Phe Val Thr Ile 305 310 315 320
- Gly Lys Ile Gly Asn Met Arg Gln Ala His Cys Asn Ile Ser Arg Ala 325 330 335
- Lys Trp Asn Asn Thr Leu Lys Gln Ile Ala Ser Lys Leu Arg Glu Gln
 340 345 350
- Phe Gly Asn Asn Lys Thr Ile Ile Phe Lys Gln Ser Ser Gly Gly Asp 355 360 365
- Pro Glu Ile Val Thr His Ser Phe Asn Cys Gly Glu Phe Phe Tyr 370 375 380
- Cys Asn Ser Thr Gln Leu Phe Asn Ser Thr Trp Phe Asn Ser Thr Trp 385 390 395 400
- Ser Thr Glu Gly Ser Asn Asn Thr Glu Gly Ser Asp Thr Ile Thr Leu 405 410 415
- Pro Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Lys Val Gly Lys 420 425 430
- Ala Met Tyr Ala Pro Pro Ile Ser Gly Gln Ile Arg Cys Ser Ser Asn 435 440 445
- Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Asn Ser Asn Asn Glu 450 460
- Ser Glu Ile Phe Arg Pro Gly Gly Gly Asp Met Arg Asp Asn Trp Arg 465 470 475 480
- Ser Glu Leu Tyr Lys Tyr Lys Val Val Lys Ile Glu Pro Leu Gly Val 485 490 495
- Ala Pro Thr Lys Ala Lys Arg Arg Val Val Gln Arg Glu Lys Arg Ala 500 505 510

Val Gly Ile Gly Ala Leu Phe Leu Gly Phe Leu Gly Ala Ala Gly Ser 515 520 525

- Thr Met Gly Ala Ala Ser Met Thr Leu Thr Val Gln Ala Arg Gln Leu 530 540
- Leu Ser Gly Ile Val Gln Gln Gln Asn Asn Leu Leu Arg Ala Ile Glu 545 550 550 560
- Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu
 565 570 575
- Gln Ala Arg Ile Leu Ala Val Glu Arg Tyr Leu Lys Asp Gln Gln Leu 580 585 590
- Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr Ala Val 595 600 605
- Pro Trp Asn Ala Ser Trp Ser Asn Lys Ser Leu Glu Gln Ile Trp Asn 610 620
- His Thr Thr Trp Met Glu Trp Asp Arg Glu Ile Asn Asn Tyr Thr Ser 625 630 635 640
- Leu Ile His Ser Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn 645 650 655
- Glu Gln Glu Leu Leu Glu Leu Asp Lys Trp Ala Ser Leu Trp Asn Trp
 660 665 670
- Phe Asn Ile Thr Asn Trp Leu Trp Tyr Ile Lys Leu Phe Ile Met Ile
 675 680 685
- Val Gly Gly Leu Val Gly Leu Arg Ile Val Phe Ala Val Leu Ser Ile 690 · 695 700
- Val Asn Arg Val Arg Gln Gly Tyr Ser Pro Leu Ser Phe Gln Thr His 705 710 715 720
- Leu Pro Thr Pro Arg Gly Pro Asp Arg Pro Glu Gly Ile Glu Glu Glu 725 730 735
- Gly Glu Arg Asp Arg Asp Arg Ser Ile Arg Leu Val Asn Gly Ser 740 745 750
- Leu Ala Leu Ile Trp Asp Asp Leu Arg Ser Leu Cys Leu Phe Ser Tyr 755 760 765
- His Arg Leu Arg Asp Leu Leu Leu Ile Val Thr Arg Ile Val Glu Leu 770 780
- Leu Gly Arg Arg Gly Trp Glu Ala Leu Lys Tyr Trp Trp Asn Leu Leu 785 790 795 800
- Gln Tyr Trp Ser Gln Glu Leu Lys Asn Ser Ala Val Ser Leu Leu Asn 805 810 815
- Ala Thr Ala Ile Ala Val Ala Glu Gly Thr Asp Arg Val Ile Glu Val 820 825 830

Val Gln Gly Ala Cys Arg Ala Ile Arg His Ile Pro Arg Arg Ile Arg 835 840 845

Gln Gly Leu Glu Arg Ile Leu Leu 850 855

<210> 2

<211> 847

<212> PRT

<213> Human immunodeficiency virus

<400> 2

Met Arg Val Lys Gly Ile Arg Lys Asn Tyr Gln His Leu Trp Arg Gly
1 10 15

Gly Thr Leu Leu Leu Gly Met Leu Met Ile Cys Ser Ala Val Glu Lys 20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Thr 35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Val

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro 65 70 75 80

Gln Glu Ile Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys 85 90 95

Asn Asn Met Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp 100 105 110

Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu 115 120 125

His Cys Thr Asn Leu Lys Asn Ala Thr Asn Thr Lys Ser Ser Asn Trp 130 135 140

Lys Glu Met Asp Arg Gly Glu Ile Lys Asn Cys Ser Phe Lys Val Thr 145 150 155 160

Thr Ser Ile Arg Asn Lys Met Gln Lys Glu Tyr Ala Leu Phe Tyr Lys 165 170 . 175

Leu Asp Val Val Pro Ile Asp Asn Asp Asn Thr Ser Tyr Lys Leu Ile 180 185 190

Asn Cys Asn Thr Ser Val Ile Thr Gln Ala Cys Pro Lys Val Ser Phe 195 200 205

Glu Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu 210 215 220

Lys Cys Asn Asp Lys Lys Phe Asn Gly Ser Gly Pro Cys Thr Asn Val 225 230 235 240

Ser Thr Val Gln Cys Thr His Gly Ile Arg Pro Val Val Ser Thr Gln 245 250 255

- Leu Leu Asn Gly Ser Leu Ala Glu Glu Gly Val Val Ile Arg Ser 260 265 270
- Glu Asn Phe Thr Asp Asn Ala Lys Thr Ile Ile Val Gln Leu Lys Glu 275 280 285
- Ser Val Glu Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser 290 295 300
- Ile Thr Ile Gly Pro Gly Arg Ala Phe Tyr Ala Thr Gly Asp Ile Ile 305 310 315 320
- Gly Asp Ile Arg Gln Ala His Cys Asn Ile Ser Gly Glu Lys Trp Asn 325 330 335
- Asn Thr Leu Lys Gln Ile Val Thr Lys Leu Gln Ala Gln Phe Gly Asn
- Lys Thr Ile Val Phe Lys Gln Ser Ser Gly Gly Asp Pro Glu Ile Val 355 360 365
- Met His Ser Phe Asn Cys Gly Gly Glu Phe Phe Tyr Cys Asn Ser Thr 370 380
- Gln Leu Phe Asn Ser Thr Trp Asn Asn Thr Ile Gly Pro Asn Asn Thr 385 390 395 400
- Asn Gly Thr Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Arg 405 410 415
- Trp Gln Glu Val Gly Lys Ala Met Tyr Ala Pro Pro Ile Arg Gly Gln
 420 425 430
- Ile Arg Cys Ser Ser Asn Ile Thr Gly Leu Leu Thr Arg Asp Gly
 435 440 445
- Gly Lys Glu Ile Ser Asn Thr Thr Glu Ile Phe Arg Pro Gly Gly 450 455 460
- Asp Met Arg Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val 465 470 475 480
- Lys Ile Glu Pro Leu Gly Val Ala Pro Thr Lys Ala Lys Arg Arg Val 485 490 495
- Val Gln Arg Glu Lys Arg Ala Val Thr Leu Gly Ala Met Phe Leu Gly 500 505 510
- Phe Leu Gly Ala Ala Gly Ser Thr Met Gly Ala Arg Ser Leu Thr Leu 515 520 525
- Thr Val Gln Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Asn 530 535 540
- Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu Gln Leu Thr 545 550 555 560

Val Trp Gly Ile Lys Gln Leu Gln Ala Arg Val Leu Ala Val Glu Arg 565 570 575

- Tyr Leu Lys Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys 580 585 590
- Leu Ile Cys Thr Thr Ala Val Pro Trp Asn Ala Ser Trp Ser Asn Lys 595 600 605
- Ser Leu Asp Gln Ile Trp Asn Asn Met Thr Trp Met Glu Trp Glu Arg 610 620
- Glu Ile Asp Asn Tyr Thr Asn Leu Ile Tyr Thr Leu Ile Glu Glu Ser 625 630 635 640
- Gln Asn Gln Gln Glu Lys Asn Glu Gln Glu Leu Leu Glu Leu Asp Lys 645 650 655
- Trp Ala Ser Leu Trp Asn Trp Phe Asp Ile Ser Lys Trp Leu Trp Tyr
 660 665 670
- Ile Lys Ile Phe Ile Met Ile Val Gly Gly Leu Val Gly Leu Arg Ile 675 680 685
- Val Phe Thr Val Leu Ser Ile Val Asn Arg Val Arg Gln Gly Tyr Ser 690 695 700
- Pro Leu Ser Phe Gln Thr Arg Phe Pro Ala Pro Arg Gly Pro Asp Arg 705 710 715 720
- Pro Glu Gly Ile Glu Glu Glu Gly Gly Glu Arg Asp Arg Ser
 725 730 735
- Ser Pro Leu Val His Gly Leu Leu Ala Leu Ile Trp Asp Asp Leu Arg
 740 745 750
- Ser Leu Cys Leu Phe Ser Tyr His Arg Leu Arg Asp Leu Ile 755 760 765
- Ala Ala Arg Ile Val Glu Leu Leu Gly Arg Arg Gly Trp Glu Ala Leu
 770 780
- Lys Tyr Trp Gly Asn Leu Leu Gln Tyr Trp Ile Gln Glu Leu Lys Asn 785 790 795 800
- Ser Ala Val Ser Leu Phe Asp Ala Ile Ala Ile Ala Val Ala Glu Gly
- Thr Asp Arg Ile Ile Glu Val Ala Gln Arg Ile Gly Arg Ala Phe Leu 820 825 830
- His Ile Pro Arg Arg Ile Arg Gln Gly Phe Glu Arg Ala Leu Leu 835 840 845

<210> 3

<211> 2310

<212> DNA

<213> Artificial Sequence

<220> <223> Description of Artificial Sequence: Vall20-Ala204 <400> 3 gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60 gcagtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120 cccgtgtgga aggaggccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180 accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240 gagategtge tggagaaegt gaeegagaae tteaacatgt ggaagaaeaa catggtggag 300 cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgggcgcc 360 ggcgcctgcc ccaaggtgag cttcgagccc atccccatcc actactgcgc ccccgccggc 420 ttcgccatcc tgaagtgcaa cgacaagaag ttcaacggca gcggcccctg caccaacgtg 480 agcaccgtgc agtgcaccca cggcatccgc cccgtggtga gcacccagct gctgctgaac 540 ggcagcctgg ccgaggaggg cgtggtgatc cgcagcgaga acttcaccga caacgccaag 600 accatcatcg tgcagctgaa ggagagcgtg gagatcaact gcacccgccc caacaacaac 660 accogcaaga qcatcaccat cggccccggc cgcgccttct acgccaccgg cgacatcatc 720 ggcgacatcc gccaggccca ctgcaacatc agcggcgaga agtggaacaa caccctgaag 780 cagategtga ccaagetgca ggcccagtte ggcaacaaga ccategtgtt caageagage 840 ageggeggeg acceegagat egtgatgeac agetteaact geggeggega gttettetac 900 tqcaacaqca cccaqctqtt caacaqcacc tggaacaaca ccatcggccc caacaacacc 960 aacggcacca tcaccctgcc ctgccgcatc aagcagatca tcaaccgctg gcaggaggtg 1020 ggcaaggcca tgtacgcccc ccccatccgc ggccagatcc gctgcagcag caacatcacc 1080 ggcctgctgc tgacccgcga cggcggcaag gagatcagca acaccaccga gatcttccgc 1140 cccggcggcg gcgacatgcg cgacaactgg cgcagcgagc tgtacaagta caaggtggtg 1200 aagatcqaqc ccctqqqcqt ggcccccacc aaggccaagc gccgcgtggt geagcgcgag 1260 aagegegeeg tgaccetggg egecatgtte etgggettee tgggegeege eggeageace 1320 atgggegece geageetgae eetgacegtg caggeeegee agetgetgag eggeategtg 1380 cagcagcaga acaacctgct gcgcgccatc gaggcccagc agcacctgct gcagctgacc 1440 gtgtggggca tcaaqcaqct gcaggccgc gtgctggccg tggagcgcta cctgaaggac 1500 cagcagetge tgggcatetg gggctgcage ggcaagetga tetgcaecae egeegtgeee 1560 tggaacgcca gctggagcaa caagagcctg gaccagatct ggaacaacat gacctggatg 1620 gagtgggagc gcgagatcga caactacacc aacctgatct acaccctgat cgaggagagc 1680 cagaaccagc aggagaagaa cgagcaggag ctgctggagc tggacaagtg ggccagcctg 1740 tggaactggt tcgacatcag caagtggctg tggtacatca agatcttcat catgatcgtg 1800 ggcggcctgg tgggcctgcg catcgtgttc accgtgctga gcatcgtgaa ccgcgtgcgc 1860 cagggetaca geoceetgag ettecagace egetteceeg ecceegegg eccegácege 1920 cccgagggca tcgaggagga ggggggag cgcgaccgcg accgcagcag ccccctggtg 1980 caeggeetge tggeeetgat etgggaegae etgegeagee tgtgeetgtt eagetaceae 2040 egectgegeg acctgatect gategeegee egeategtgg agetgetggg eegeegegge 2100 tgggaggccc tgaagtactg gggcaacctg ctgcagtact ggatccagga gctgaagaac 2160 agogoogtga gootgttoga ogocatogoo atogoogtgg oogagggcao ogacogoato 2220 ategaggtgg cocagegeat eggeegegee tteetgeaca tecceegeeg cateegeeag 2280 ggcttcgagc gcgccctgct gtaactcgag 2310 <210> 4 <211> 2316 <212> DNA <213> Artificial Sequence <220> <223> Description of Artificial Sequence: Vall20-Ile201 gaattegeea ceatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60 gragitating thingcocag cgccgtggag aagctgtggg tgaccgtgta chacggcgtg 120 cccgtgtgga aggaggccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180 accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240 gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300 cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgggcggc 360

```
atcacccagg cctgccccaa ggtgagcttc gagcccatcc ccatccacta ctgcgccccc 420
gccggcttcg ccatcctgaa gtgcaacgac aagaagttca acggcagcgg cccctgcacc 480
aacgtgagca cogtgcagtg cacccacggc atccgccccg tggtgagcac ccaqctqctq 540
ctgaacggca gcctggccga ggagggcgtg gtgatccgca gcgagaactt caccgacaac 600
gccaagacca tcatcgtgca gctgaaggag agcgtggaga tcaactgcac ccgccccaac 660
aacaacaccc gcaagagcat caccatcggc cccggccgcg ccttctacgc caccggcgac 720
atcatcggcg acatccgcca ggcccactgc aacatcagcg gcgagaagtg gaacaacacc 780
ctgaagcaga tcgtgaccaa gctgcaggcc cagttcggca acaagaccat cgtgttcaag 840
cagagcagcg geggegaccc egagategtg atgeacaget teaactgegg eggegagtte 900
ttctactgca acagcaccca gctgttcaac agcacctgga acaacaccat cgqccccaac 960
aacaccaacg gcaccatcac cctgccctgc cgcatcaagc agatcatcaa ccgctggcag 1020
gaggtgggca aggccatgta cgccccccc atccgcggcc agatccgctg cagcagcaac 1080
atcaccggcc tgctgctgac ccgcgacggc ggcaaggaga tcagcaacac caccgagatc 1140
ttccgccccg gcggcggcga catgcgcgac aactggcgca gcgagctgta caagtacaag 1200
gtggtgaaga tcgagcccct gggcgtggcc cccaccaagg ccaagcgccg cgtggtgcag 1260
cgcgagaagc gcgccgtgac cctgggcgcc atgttcctgg gcttcctggg cgccgccggc 1320
agcaccatgg gcgcccgcag cctgaccctg accgtgcagg cccgccagct gctgagcggc 1380
atogtgcagc agcagaacaa cotgotgcgc gccatcgagg cocagcagca cotgotgcag 1440
ctgaccgtgt ggggcatcaa gcagctgcag gcccgcgtgc tggccgtgga gcgctacctg 1500
aaggaccagc agctgctggg catctggggc tgcagcggca agctgatctg caccaccqcc 1560
gtgccctgga acgccagctg gagcaacaag agcctggacc aqatctqqaa caacatqacc 1620
tggatggagt gggagcgca gatcgacaac tacaccaacc tgatctacac cctgatcgag 1680
gagagccaga accagcagga gaagaacgag caggagctgc tggagctgga caagtgggcc 1740
agcetgtgga actggttega cateageaag tggetgtggt acateaagat etteateatg 1800
atogtgggcg gcctggtggg cctgcgcatc gtgttcaccg tgctgagcat cgtgaaccgc 1860
gtgcgccagg gctacagccc cctgagcttc cagaccegct tccccgcccc ccqcqqcccc 1920
gaccgccccg agggcatcga ggaggaggc ggcgagcgcg accgcgaccg cagcagcccc 1980
ctggtgcacg gcctgctggc cctgatctgg gacgacctgc gcagcctgtg cctgttcagc 2040
taccaccgcc tgcgcgacct gatcctgatc gccgcccgca tcgtgqagct qctqqqccqc 2100
cgcggctggg aggccctgaa gtactggggc aacctgctgc agtactggat ccaggagctg 2160
aagaacagcg ccgtgagcct gttcgacgcc atcgccatcg ccgtggccga gggcaccgac 2220
cgcatcatcg aggtggccca gcgcatcggc cgcgccttcc tgcacatccc ccgccgcatc 2280
cgccagggct tcgagcgcgc cctgctgtaa ctcgag
<210> 5
<211> 2322
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Val120-Ile201B
<400> 5
gaattegeea ceatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60
gcagtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccgtgtgga aggaggccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccqaccc caaccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgcccggc 360
atcacccagg cotgocccaa ggtgagotto gagoccatoo coatcoacta otgogoccoo 420
gccggcttcg ccatcctgaa gtgcaacgac aagaagttca acggcagcgg cccctgcacc 480
aacgtgagca cogtgcagtg cacccaoggc atcogccccg tggtgagcac coagctgctg 540
ctgaacggca gcctggccga ggagggcgtg gtgatccgca gcgagaactt caccgacaac 600
gccaagacca tcatcgtgca gctgaaggag agcgtggaga tcaactgcac ccgccccaac 660
aacaacacc gcaagagcat caccategge eceggeegeg cettetaege caceggegae 720
atcateggeg acateegeea ggeecaetge aacateageg gegagaagtg gaacaacace 780
ctgaagcaga tegtgaccaa getgeaggee cagtteggea acaagaceat egtgtteaag 840
cagagcageg geggegacee egagategtg atgeacaget teaactgegg eggegagtte 900
```

ttctactgca acagcaccca gctgttcaac agcacctgga acaacaccat cggccccaac 960

```
aacaccaacg gcaccatcac cctgccctgc cgcatcaagc agatcatcaa ccgctggcag 1020
gaggtgggca aggccatgta cgccccccc atccgcggcc agatccgctg cagcagcaac 1080
atcaccacc tactactac cogcacaca agraagaaa tcagcaacac caccaagate 1140
ttccgccccg gcggcggcga catgcgcgac aactggcgca gcgagctgta caagtacaag 1200
qtqqtqaaqa tcgagcccct gggcgtggcc cccaccaagg ccaagcgccg cgtggtgcag 1260
cgcgagaage gegeegtgae eetgggegee atgtteetgg getteetggg egeegeegge 1320
agcaccatgg gegecegeag cetgaecetg accgtgeagg ecegecaget getgagegge 1380
atcqtqcaqc aqcagaacaa cctgctgcgc gccatcgagg cccagcagca cctgctgcag 1440
ctgaccgtgt ggggcatcaa gcagctgcag gcccgcgtgc tggccgtgga gcgctacctg 1500
aaggaccagc agctgctggg catctggggc tgcagcggca agctgatctg caccaccgcc 1560
gtgccctgga acgccagctg gagcaacaag agcctggacc agatctggaa caacatgacc 1620
tggatggagt gggagcgcga gatcgacaac tacaccaacc tgatctacac cctgatcgag 1680
gagagccaga accagcagga gaagaacgag caggagctgc tggagctgga caagtgggcc 1740
agectgtgga actggttcga catcagcaag tggctgtggt acatcaagat cttcatcatg 1800
atcgtgggcg gcctggtggg cctgcgcatc gtgttcaccg tgctgagcat cgtgaaccgc 1860
gtgcgccagg gctacagccc cctgagcttc cagacccgct tccccgcccc ccgcggcccc 1920
gaccgccccg agggcatcga ggaggagggc ggcgagcgcg accgcgaccg cagcagcccc 1980
ctggtgcacg gcctgctggc cctgatctgg gacgacctgc gcagcctgtg cctgttcagc 2040
taccaccgcc tgcgcgacct gatcctgatc gccgcccgca tcgtggagct gctgggccgc 2100
eqeqqetqqq aqqeeetgaa gtactgggge aacetgetge agtactggat ecaggagetg 2160
aagaacageg cegtgageet gttegaegee ategecateg cegtggeega gggeaecgae 2220
cgcatcatcg aggtggccca gcgcatcggc cgcgccttcc tgcacatccc ccgccgcatc 2280
cgccagggct tcgagcgcgc cctgctgtaa ctcgagcgtg ct
<210> 6
<211> 2328
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Lys121-Val200
<400> 6
gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
gcagtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccgtgtgga aggaggccac caccaccctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagategtge tggagaaegt gaeegagaae tteaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaaggcc 360
cccgtgatca cccaggcctg ccccaaggtg agcttcgagc ccatccccat ccactactgc 420
gcccccgccg gcttcgccat cctgaagtgc aacgacaaga agttcaacgg cagcgcccc 480
tgcaccaacg tgagcaccgt gcagtgcacc cacggcatcc gccccgtggt gagcacccag 540
ctgctgctga acggcagcct ggccgaggag ggcgtggtga tccgcagcga gaacttcacc 600
gacaacgcca agaccatcat cgtgcagctg aaggagagcg tggagatcaa ctgcacccgc 660
cccaacaaca acaccegcaa gagcatcacc ateggecceg geogegeett ctacgecacc 720
ggcgacatca tcggcgacat ccgccaggcc cactgcaaca tcagcggcga gaagtggaac 780
aacaccetga agcagategt gaccaagetg caggeccagt teggeaacaa gaccategtg 840
ttcaagcaga gcagcggcgg cgaccccgag atcgtgatgc acagcttcaa ctgcggcggc 900
gagttettet actgeaacag cacceagetg tteaacagea cetggaacaa caccategge 960
cccaacaaca ccaacggcac catcaccctg ccctgccgca tcaagcagat catcaaccgc 1020
tggcaggagg tgggcaaggc catgtacgcc cccccatcc gcggccagat ccgctgcagc 1080
agcaacatca coggootgot gotgaccogo gaoggoggoa aggagatcag caacaccaco 1140
gagatettee geeeeggegg eggegacatg egegacaaet ggegeagega getgtacaag 1200
tacaaggtgg tgaagatcga gcccctgggc gtggcccca ccaaggccaa gcgccgcgtg 1260
gtgcagcgcg agaagcgcgc cgtgaccctg ggcgccatgt tcctgggctt cctgggcgcc 1320
geoggeagea ceatgggege cegeageetg accetgaceg tgeaggeeeg ceagetgetg 1380
ageggeateg tgeageagea gaacaacetg etgegegeea tegaggeeea geageacetg 1440
ctgcagctga ccgtgtgggg catcaagcag ctgcaggccc gcgtgctggc cgtggagcgc 1500
tacctgaagg accagcagct gctgggcatc tggggctgca gcggcaagct gatctgcacc 1560
```

```
accgccgtgc cctggaacgc cagctggagc aacaagagcc tggaccagat ctggaacaac 1620
atgacctgga tggagtggga gcgcgagatc gacaactaca ccaacctgat ctacaccctg 1680
atcgaggaga gccagaacca gcaggagaag aacgagcagg agctgctgga gctggacaag 1740
tgggccagcc tgtggaactg gttcgacatc agcaagtggc tgtggtacat caagatcttc 1800
atcatgatcg tgggcggcct ggtgggcctg cgcatcgtgt tcaccgtgct gagcatcgtg 1860
aaccgcgtgc gccagggcta cagcccctg agcttccaga cccgcttccc cgcccccgc 1920
ggccccgacc gccccgaggg catcgaggag gagggcggcg agcgcgaccg cgaccgcagc 1980
agececting tigeaeggeet getiggeeetg atetinggaeg acctingeag cetinities 2040
ttcagctacc accgectgeg egacetgate etgategeeg ecegeategt ggagetgetg 2100
ggccgccgcg gctgggaggc cctgaagtac tggggcaacc tgctgcagta ctggatccag 2160
gagetgaaga acagegeegt gageetgtte gaegeeateg ceategeegt ggeegaggge 2220
accgaccgca tcatcgaggt ggcccagcgc atcggccgcg ccttcctgca catcccccgc 2280
cgcatccgcc agggcttcga gcgcgccctg ctgtaactcg agcgtgct
<210> 7
<211> 2334
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Leu122-Ser199
<400> 7
gaattegeea ceatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60
gcagtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccgtgtgga aggaggccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagategtge tggagaaegt gacegagaae tteaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360
ggcaacagcg tgatcaccca ggcctgcccc aaggtgagct tcgagcccat ccccatccac 420
tactgcgccc ccgccggctt cgccatcctg aagtgcaacg acaagaagtt caacggcagc 480
ggcccctgca ccaacgtgag caccgtgcag tgcacccacg gcatccgccc cgtggtgagc 540
acceagetge tgetgaaegg cageetggee gaggagggeg tggtgateeg cagegagaac 600
ttcaccgaca acgccaagac catcatcgtg cagctgaagg agagcgtgga gatcaactgc 660
accegececa acaacaacae eegcaagage ateaecateg geeeeggeeg egeettetae 720
gccaccggcg acatcatcgg cgacatccgc caggcccact gcaacatcag cggcgagaag 780
tggaacaaca coctgaagca gategtgacc aagctgcagg cocagttegg caacaagacc 840
atogtgttca agcagagcag oggoggogac ocogagatog tgatgcacag ottcaactgo 900
ggcggcgagt tettetactg caacagcace cagetgttea acageacetg gaacaacace 960
atoggococa acaacaccaa oggoaccato accotgocot googcatcaa goagatcato 1020
aaccgctggc aggaggtggg caaggccatg tacgccccc ccatccgcgg ccagatccgc 1080
tgcagcagca acatcaccgg cctgctgctg acccgcgacg gcggcaagga gatcagcaac 1140
accacegaga tetteegeee eggeggegge gacatgegeg acaactggeg cagegagetg 1200
tacaagtaca aggtggtgaa gatcgagccc ctgggcgtgg cccccaccaa ggccaagcgc 1260
cgcgtggtgc agcgcgagaa gcgcgccgtg accctgggcg ccatgttcct gggcttcctg 1320
ggcgccgccg gcagcaccat gggcgcccgc agcctgaccc tgaccgtgca ggcccgccag 1380
ctgctgagcg gcatcgtgca gcagcagaac aacctgctgc gcgccatcga ggcccagcag 1440
cacctgctgc agctgaccgt gtggggcatc aagcagctgc aggcccgcgt gctggccgtg 1500
gagegetace tgaaggacea geagetgetg ggeatetggg getgeagegg caagetgate 1560
tgcaccaccg ccgtgccctg gaacgccagc tggagcaaca agagcctgga ccagatctgg 1620
aacaacatga cctggatgga gtgggagcgc gagatcgaca actacaccaa cctgatctac 1680
accetgateg aggagageea gaaccageag gagaagaaeg ageaggaget getggagetg 1740
gacaagtggg ccagcctgtg gaactggttc gacatcagca agtggctgtg gtacatcaag 1800
atcttcatca tgatcgtggg cggcctggtg ggcctgcgca tcgtgttcac cgtgctgagc 1860
ategtgaacc gegtgegeca gggetacage eccetgaget tecagacceg etteecegee 1920
ccccgcggcc ccgaccgccc cgagggcatc gaggaggagg gcggcgagcg cgaccgcgac 1980
cgcagcagcc ccctggtgca cggcctgctg gccctgatct gggacgacct gcgcagcctg 2040
tgcctgttca gctaccaccg cctgcgcgac ctgatcctga tcgccgcccg catcgtggag 2100
ctgctgggcc gccgcggctg ggaggccctg aagtactggg gcaacctgct gcagtactgg 2160
```

```
atccaggage tgaagaacag egeegtgage etgttegaeg ceategeeat egeegtggee 2220
gagggcaccg accgcatcat cgaggtggcc cagcgcatcg gccgcgcctt cctgcacatc 2280
ccccgccqca tccgccaggg cttcgagcgc gccctgctgt aactcgagcg tgct
<210> 8
<211> 2316
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Vall20-Thr202
gaattegeca ccatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60
geagtetteg tttegeceag egeegtggag aagetgtggg tgacegtgta etaeggegtg 120
cccqtqtqqa aqqaqqccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgggcggc 360
gccacccagg cctgccccaa ggtgagcttc gagcccatcc ccatccacta ctgcgccccc 420
qeeqqetteq ceatectqaa gtgcaacgac aagaagttca acggcagegg cecetgcace 480
aacqtgaqca ccgtgcagtg cacccacggc atccgccccg tggtgagcac ccagctgctg 540
ctgaacggca gcctggccga ggagggcgtg gtgatccgca gcgagaactt caccgacaac 600
qccaaqacca tcatcgtgca gctgaaggag agcgtggaga tcaactgcac ccgccccaac 660
aacaacaccc qcaaqaqcat caccatcggc cccggccgcg ccttctacgc caccggcgac 720
atcatcggcg acatccgcca ggcccactgc aacatcagcg gcgagaagtg gaacaacacc 780
ctgaagcaga tcgtgaccaa gctgcaggcc cagttcggca acaagaccat cgtgttcaag 840
cagageageg geggegacce egagategtg atgeacaget teaactgegg eggegagtte 900
ttctactqca acaqcaccca qctqttcaac agcacctqqa acaacaccat cgqccccaac 960
aacaccaacg qcaccatcac cctgccctgc cgcatcaagc agatcatcaa ccgctggcag 1020
gaggtgggca aggccatgta cgccccccc atccgcggcc agatccgctg cagcagcaac 1080
atcaccggcc tgctgctgac ccgcgacggc ggcaaggaga tcagcaacac caccgagatc 1140
ttccgccccg gcggcggcga catgcgcgac aactggcgca gcgagctgta caagtacaag 1200
gtggtgaaga tcgagcccct gggcgtggcc cccaccaagg ccaagcgccg cgtggtgcag 1260
cgcgagaagc gcgccgtgac cctgggcgcc atgttcctgg gcttcctggg cgccgccggc 1320
agcaccatgg gegecegeag cetgaccetg accgtgeagg ceegecaget getgagegge 1380
ategtgcage ageagaacaa cetgetgege gecategagg cecageagea cetgetgcag 1440
ctgaccgtgt ggggcatcaa gcagctgcag gcccgcgtgc tggccgtgga gcgctacctg 1500
aaggaccagc agetgetggg catetgggge tgeageggea agetgatetg caccacegee 1560
gtgccctgga acgccagctg gagcaacaag agcctggacc agatctggaa caacatgacc 1620
tggatggagt gggagcgcga gatcgacaac tacaccaacc tgatctacac cctgatcgag 1680
qaqaqccaqa accaqcagga gaagaacgag caggagctgc tggagctgga caagtgggcc 1740
agcctgtgga actggttcga catcagcaag tggctgtggt acatcaagat cttcatcatg 1800
atcgtgggcg gcctggtggg cctgcgcatc gtgttcaccg tgctgagcat cgtgaaccgc 1860
gtgcgccagg gctacagccc cctgagcttc cagacccgct tccccgcccc ccgcggcccc 1920
gaccgccccg agggcatcga ggaggaggc ggcgagcgcg accgcgaccg cagcagcccc 1980
ctgqtqcacq qcctqctqqc cctgatctgg gacgacctgc gcagcctgtg cctgttcagc 2040
taccaccgcc tgcgcgacct gatectgate gccgcccgca tcgtggagct gctgggccgc 2100
cgcggctggg aggccctgaa gtactggggc aacctgctgc agtactggat ccaggagctg 2160
aagaacagcg ccgtgagcct gttcgacgcc atcgccatcg ccgtggccga gggcaccgac 2220
cgcatcatcg aggtggccca gcgcatcggc cgcgccttcc tgcacatccc ccgccgcatc 2280
cgccagggct tcgagcgcgc cctgctgtaa ctcgag
                                                                  2316
<210> 9
<211> 2541
<212> DNA
<213> Artificial Sequence
<220>
```

<223> Description of Artificial Sequence: Trp427-Gly431

```
qaattcqcca ccatqqatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
gcagtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccgtgtgga aggaggccac caccaccctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagategtge tggagaacgt gacegagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360
accccctgt gcgtgaccct gcactgcacc aacctgaaga acgccaccaa caccaagagc 420
agcaactgga aggagatgga ccgcggcgag atcaagaact gcagcttcaa ggtgaccacc 480
agcatccqca acaagatgca gaaggagtac gccctgttct acaagctgga cgtggtgccc 540
atcgacaacg acaacaccag ctacaagctg atcaactgca acaccagcgt gatcacccag 600
gcctgcccca aggtgagett cgageccate eccatecaet actgegeece egeeggette 660
qccatcctqa agtgcaacga caagaagttc aacggcagcg gcccctgcac caacgtgagc 720
accgtgcagt gcacccacgg catccgcccc gtggtgagca cccagctgct gctgaacggc 780
agcctggccg aggagggcgt ggtgatccgc agcgagaact tcaccgacaa cgccaagacc 840
atcatcgtgc agctgaagga gagcgtggag atcaactgca cccgccccaa caacaacacc 900
cqcaaqaqca tcaccatcgg ccccggccgc gccttctacg ccaccggcga catcatcggc 960
gacatccgcc aggcccactg caacatcagc ggcgagaagt ggaacaacac cctgaagcag 1020
atcgtgacca agctgcaggc ccagttcggc aacaagacca tcgtgttcaa gcagagcagc 1080
ggcggcgacc ccgagatcgt gatgcacagc ttcaactgcg gcggcgagtt cttctactgc 1140
aacagcaccc agctgttcaa cagcacctgg aacaacacca toggccccaa caacaccaac 1200
ggcaccatca ccctgccctg ccgcatcaag cagatcatca accgctgggg cggcaaggcc 1260
atgtacgccc cccccatccg cggccagatc cgctgcagca gcaacatcac cggcctgctg 1320
ctgacccgcg acggcggcaa ggagatcagc aacaccaccg agatcttccg ccccggcggc 1380
ggcgacatgc gcgacaactg gcgcagcgag ctgtacaagt acaaggtggt gaagatcgag 1440
cccctgggcg tggcccccac caaggccaag cgccgcgtgg tgcagcgcga gaagcgcgcc 1500
gtgaccctgg gcgccatgtt cctgggcttc ctgggcgccg ccggcagcac catgggcgcc 1560
cgcagcctga ccctgaccgt gcaggcccgc cagctgctga gcggcatcgt gcagcagcag 1620
aacaacetge tgegegeeat egaggeeeag cageacetge tgeagetgae egtgtgggge 1680
atcaagcagc tgcaggcccg cgtgctggcc gtggagcgct acctgaagga ccagcagctg 1740
ctgggcatct ggggctgcag cggcaagctg atctgcacca ccgccgtgcc ctggaacgcc 1800
agctggagca acaagagcct ggaccagatc tggaacaaca tgacctggat ggagtgggag 1860
cgcgagatcg acaactacac caacctgatc tacaccctga tcgaggagag ccagaaccag 1920
caggagaaga acgagcagga gctgctggag ctggacaagt gggccagcct gtggaactgg 1980
ttcgacatca gcaagtggct gtggtacatc aagatcttca tcatgatcgt gggcggcctg 2040
gtgggcctgc gcatcgtgtt caccgtgctg agcatcgtga accgcgtgcg ccagggctac 2100
agecceetga gettecagae cegetteece gecceeegeg geccegaeeg eecegaggge 2160
atcgaggagg agggcggcga gcgcgaccgc gaccgcagca gccccctggt gcacggcctg 2220
ctqqccctqa tctqqqacga cctgcgcagc ctgtgcctgt tcagctacca ccgcctgcgc 2280
gacctgatee tgategeege cegeategtg gagetgetgg geegeegegg etgggaggee 2340
ctgaagtact ggggcaacct gctgcagtac tggatccagg agctgaagaa cagcgccgtg 2400
agectgtteg aegecatege categeegtg geegagggea eegacegeat categaggtg 2460
geocagegea teggeegege etteetgeac atecceegee geateegeea gggettegag 2520
cgcgccctgc tgtaactcga g
<210> 10
<211> 2541
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Arg426-Gly431
<400> 10
gaattegeea ceatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60
gcagtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
```

cccgtgtgga aggaggccac caccaccctg ttctgcgcca gcgacgccaa ggcctacgac 180

WO 00/39303 PCT/US99/31272 accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caaccccag 240

gagategtge tggagaaegt gacegagaae tteaacatgt ggaagaacaa catggtggag 300

```
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360
acceccetgt gegtgaccet geactgeace aacetgaaga acgecaceaa caccaagage 420
agcaactgga aggagatgga ccgcggcgag atcaagaact gcagcttcaa ggtgaccacc 480
agcatcegea acaagatgea gaaggagtae geeetgttet acaagetgga egtggtgeee 540
ategacaaeg acaacaceag etacaagetg ateaaetgea acaceagegt gateaeceag 600
geotgeocca aggtgagett egageocate cecatecaet actgegeece egeoggette 660
gccatcctga agtgcaacga caagaagttc aacggcagcg gcccctgcac caacgtgagc 720
acceptgcagt gcacccacge catcceccc gtggtgagca cccagctgct gctgaacggc 780
agectggeeg aggagggegt ggtgateege agegagaaet teaeegaeaa egeeaagaee 840
atcategtge agetgaagga gagegtggag atcaactgea eeegeeecaa caacaacaee 900
cgcaagagca tcaccatcgg ccccggccgc gccttctacg ccaccggcga catcatcggc 960
gacatcegee aggeecactg caacateage ggegagaagt ggaacaacae cetgaageag 1020
atogtgacca agotgoaggo coagttoggo aacaagacca togtgttoaa goagagoago 1080
ggcggcgacc ccgagatcgt gatgcacagc ttcaactgcg gcggcgagtt cttctactgc 1140
aacagcaccc agctgttcaa cagcacctgg aacaaccac tcggccccaa caacaccaac 1200
ggcaccatca ccctgccctg ccgcatcaag cagatcatca accgcggcgg cggcaaggcc 1260
atgtacgccc ccccatccg cggccagatc cgctgcagca gcaacatcac cggcctgctg 1320
ctgaccogcg acggcggcaa ggagatcagc aacaccaccg agatcttccg ccccggcggc 1380
ggcgacatgc gcgacaactg gcgcagcgag ctgtacaagt acaaggtggt gaagatcgag 1440
cccctgggcg tggccccac caaggccaag cgccgcgtgg tgcagcgcga gaagcgcgcc 1500
gtgaccotgg gcgccatgtt cotgggctto ctgggcgccg ccggcagcac catgggcgcc 1560
cgcagcctga ccctgaccgt gcaggcccgc cagctgctga gcggcatcgt gcagcagcag 1620
aacaacetge tgegegeeat egaggeeeag eageacetge tgeagetgae egtgtgggge 1680
atcaagcagc tgcaggcccg cgtgctggcc gtggagcgct acctgaagga ccagcagctg 1740
ctgggcatct ggggctgcag cggcaagctg atctgcacca ccgccgtgcc ctggaacgcc 1800
agctggagca acaagagcct ggaccagatc tggaacaaca tgacctggat ggagtgggag 1860
cgcgagatcg acaactacac caacctgatc tacaccctga tcgaggagag ccagaaccag 1920
caggagaaga acgagcagga gctgctggag ctggacaagt gggccagcct gtggaactgg 1980
ttcgacatca gcaagtggct gtggtacatc aagatcttca tcatgatcgt gggcggcctg 2040
gtgggcctgc gcatcgtgtt caccgtgctg agcatcgtga accgcgtgcg ccagggctac 2100
agececetga gettecagae eegetteeee geeeeegeg geeeegaeeg eeeegaggge 2160
atcgaggagg agggcggcga gcgcgaccgc gaccgcagca gccccctggt gcacggcctg 2220
ctggccctga tctgggacga cctgcgcagc ctgtgcctgt tcagctacca ccgcctgcgc 2280
gacctgatec tgategeege eegcategtg gagetgetgg geegeegegg etgggaggee 2340
ctgaagtact ggggcaacct gctgcagtac tggatccagg agctgaagaa cagcgccgtg 2400
agcctgttcg acgccatcgc catcgccgtg gccgagggca ccgaccgcat catcgaggtg 2460
geccagegea teggeegege etteetgeae atecceegee geateegeea gggettegag 2520
cgcgccctgc tgtaactcga g
<210> 11
<211> 2541
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Arg426-Gly431B
<400> 11
gaattegeea ceatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60
gcagtetteg tttegeccag egcegtggag aagetgtggg tgacegtgta etaeggegtg 120
cccgtgtgga aggaggccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360
acceccetgt gegtgaccet geactgeace aacetgaaga aegecaceaa caccaagage 420
agcaactgga aggagatgga ccgcggcgag atcaagaact gcagcttcaa ggtgaccacc 480
agcatcogca acaagatgca gaaggagtac gccctgttct acaagctgga cgtggtgccc 540
```

```
atogacaacg acaacaccag ctacaagctg atcaactgca acaccagcgt gatcacccag 600
geotgeecca aggtgagett egageecate eccatecaet actgegeece egeeggette 660
qccatcctga agtgcaacga caagaagttc aacggcagcg gcccctgcac caacgtgagc 720
acceptgcagt gcacccacgg catccgcccc gtggtgagca cccagctgct gctgaacggc 780
agectggeeg aggagggegt ggtgateege agegagaact teacegaeaa egeeaagaee 840
atcatcgtgc agctgaagga gagcgtggag atcaactgca cccgccccaa caacaacacc 900
cqcaaqagca tcaccatcgg ccccggccgc gccttctacg ccaccggcga catcatcggc 960
qacatccqcc aggcccactg caacatcage ggcgagaagt ggaacaacac cctgaagcag 1020
atogtgacca agotgcaggo coagttoggo aacaagacca togtgttoaa goagagcago 1080
ggcggcgacc ccgagatcgt gatgcacagc ttcaactgcg gcggcgagtt cttctactgc 1140
aacagcaccc agctgttcaa cagcacctgg aacaacacca tcggccccaa caacaccaac 1200
ggcaccatca ccctgccctg ccgcatcaag cagatcatca accgcggcag cggcaaggcc 1260
atgtacgccc cccccatccg cggccagatc cgctgcagca gcaacatcac cggcctgctg 1320
ctgacccgcg acggcggcaa ggagatcagc aacaccaccg agatcttccg ccccggcggc 1380
ggcgacatgc gcgacaactg gcgcagcgag ctgtacaagt acaaggtggt gaagatcgag 1440
cccctgggcg tggcccccac caaggccaag cgccgcgtgg tgcagcgcga gaagcgcgcc 1500
gtgaccctgg gcgccatgtt cctgggcttc ctgggcgccg ccggcagcac catgggcgcc 1560
cgcagcctga ccctgaccgt gcaggcccgc cagctgctga gcggcatcgt gcagcagcag 1620
aacaacctqc tqcgcgccat cgaggcccag cagcacctgc tgcagctgac cgtgtggggc 1680
atcaaqcaqc tqcaggcccg cgtgctggcc gtggagcgct acctgaagga ccagcagctg 1740
ctgggcatct ggggctgcag cggcaagctg atctgcacca ccgccgtgcc ctggaacgcc 1800
agctggagca acaagagcct ggaccagatc tggaacaaca tgacctggat ggagtgggag 1860
cgcgagatcg acaactacac caacctgatc tacaccctga tcgaggagag ccagaaccag 1920
caqqaqaaqa acqaqcagga gctgctggag ctggacaagt gggccagcct gtggaactgg 1980
ttcgacatca gcaagtggct gtggtacatc aagatcttca tcatgatcgt gggcggcctg 2040
gtgggcctgc gcatcgtgtt caccgtgctg agcatcgtga accgcgtgcg ccagggctac 2100
ageccectga gettecagae eegetteeee geeeeegeg geeeegaeeg eeeegaggge 2160
atcgaggagg agggcggcga gcgcgaccgc gaccgcagca gcccctggt gcacggcctg 2220
ctggccctga tctgggacga cctgcgcagc ctgtgcctgt tcagctacca ccgcctgcgc 2280
gacctgatcc tgatcgccgc ccgcatcgtg gagctgctgg gccgccgcgg ctgggaggcc 2340
ctgaagtact ggggcaacct gctgcagtac tggatccagg agctgaagaa cagcgccgtg 2400
agcetgtteg acgecatege categeegtg geegagggea eegacegeat categaggtg 2460
qcccaqcqca tcqqccgcgc cttcctgcac atcccccgcc gcatccgcca gggcttcgag 2520
cgcgccctgc tgtaactcga g
                                                                  2541
<210> 12
<211> 2541
<212> DNA
<213> Artificial Sequence
<223> Description of Artificial Sequence: Arg426-Lys432
<400> 12
gaattegeea ceatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60
gcagtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccgtgtgga aggaggccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtge acaacgtgtg ggccacccac gcctgcgtge ccaccgaccc caacccccag 240
gagategtge tggagaaegt gacegagaae tteaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360
accecetqt geqtgaceet geactgeace aacetgaaga acgeeaceaa caccaagage 420
agcaactgga aggagatgga ccgcggcgag atcaagaact gcagcttcaa ggtgaccacc 480
agcatccgca acaagatgca gaaggagtac gccctgttct acaagctgga cgtggtgccc 540
ategacaacg acaacaccag ctacaagctg atcaactgca acaccagcgt gatcacccag 600
gcctgcccca aggtgagctt cgagcccatc cccatccact actgcgcccc cgccggcttc 660
gccatcctga agtgcaacga caagaagttc aacggcagcg gcccctgcac caacgtgagc 720
acceptgcagt gcacccacge catcceccce gtggtgagca cccagctgct gctgaacggc 780
agcctqgccq aggagggcgt ggtgatccgc agcgagaact tcaccgacaa cgccaagacc 840
```

atcategtge agetgaagga gagegtggag atcaactgea ceegeceeaa caacaacace 900

```
cgcaagagca tcaccatcgg ccccggccgc gccttctacg ccaccggcga catcatcggc 960
qacatccqcc aggcccactg caacatcagc ggcgagaagt ggaacaacac cctgaagcag 1020
ategtgaeca agetgeagge ceagttegge aacaagaeca tegtgtteaa geagageage 1080
ggcggcgacc ccgagatcgt gatgcacagc ttcaactgcg gcggcgagtt cttctactgc 1140
aacagcaccc agctgttcaa cagcacctgg aacaacacca tcggccccaa caacaccaac 1200
ggcaccatca ccctgccctg ccgcatcaag cagatcatca accgcggcgg caacaaggcc 1260
atgtacgece eccecateeg eggecagate egetgeagea geaacateae eggeetgetg 1320
ctgacccgcg acggcggcaa ggagatcagc aacaccaccg agatettecg ccccggcggc 1380
ggcgacatgc gcgacaactg gcgcagcgag ctgtacaagt acaaggtggt gaagatcgag 1440
cccctgggcg tggcccccac caaggccaag cgccgcgtgg tgcagcgcga gaagcgcgcc 1500
gtgaccetgg gegecatgtt cetgggette etgggegeeg eeggeageac catgggegee 1560
cgcagcctga ccctgaccgt gcaggcccgc cagctgctga gcggcatcgt gcagcagcag 1620
aacaacctgc tgcgcgccat cgaggcccag cagcacctgc tgcagctgac cgtgtggggc 1680
atcaagcagc tgcaggcccg cgtgctggcc gtggagcgct acctgaagga ccagcagctg 1740
ctgggcatct ggggctgcag cggcaagctg atctgcacca ccgccgtgcc ctggaacgcc 1800
agctggagca acaagagcct ggaccagatc tggaacaaca tgacctggat ggagtgggag 1860
cgcgagatcg acaactacac caacctgatc tacaccctga tcgaggagag ccagaaccag 1920
caggagaaga acgagcagga gctgctggag ctggacaagt gggccagcct gtggaactgg 1980
ttegacatea geaagtgget gtggtacate aagatettea teatgategt gggeggeetg 2040
gtqqqcctqc qcatcqtqtt caccqtqctg agcatcqtga accqcqtqcg ccaggqctac 2100
ageccectga gettecagae cegettecee geccecegeg geccegaceg cecegaggge 2160
ategaggagg agggeggega gegegaeege gaeegeagea geeeeetggt geaeggeetg 2220
ctggccctga tctgggacga cctgcgcagc ctgtgcctgt tcagctacca ccgcctgcgc 2280
qacctqatec tqateqeeqe eegcategtg gagetgetgg geegeegegg etgggaggee 2340
ctgaagtact ggggcaacct gctgcagtac tggatccagg agctgaagaa cagcgccgtg 2400
agectgtteg aegecatege categoegtg geegagggca cegacegeat categaggtg 2460
geccagegea teggeegege etteetgeae atecceegee geateegeea gggettegag 2520
cgcgccctgc tgtaactcga g
<210> 13
<211> 2535
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Asn425-Lys432
<400> 13
gaattegeea ceatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60
gcagtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccgtgtgga aggaggccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagategtge tggagaaegt gaeegagaae tteaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360
accccctgt gcgtgaccct gcactgcacc aacctgaaga acgccaccaa caccaagagc 420
agcaactgga aggagatgga ccgcggcgag atcaagaact gcagcttcaa ggtgaccacc 480
agcatccgca acaagatgca gaaggagtac gccctgttct acaagctgga cgtggtgccc 540
ategacaaeg acaacaecag etacaagetg ateaaetgea acaecagegt gateaeceag 600
geotgeocca aggtgagett egageocate eccatecaet aetgegeece egeoggette 660
gccatcctga agtqcaacga caagaagttc aacggcagcg gcccctgcac caacgtgagc 720
acceptgcagt gcacccacgg catcegeeee gtggtgagca eccagetget getgaacgge 780
agcetggeeg aggagggegt ggtgateege agegagaaet teacegaeaa egeeaagaee 840
atcategtge agetgaagga gagegtggag ateaaetgea eeegeeceaa caacaacaee 900
cgcaagagca tcaccategg ccceggccgc gccttctacg ccaccggcga catcateggc 960
gacatcogec aggeccactg caacatcage ggegagaagt ggaacaacac cetgaagcag 1020
atcgtgacca agctgcaggc ccagttcggc aacaagacca tcgtgttcaa gcagagcagc 1080
ggcggcgacc ccgagatcgt gatgcacagc ttcaactgcg gcggcgagtt cttctactgc 1140
aacagcaccc agctgttcaa cagcacctgg aacaacacca tcggccccaa caacaccaac 1200
ggcaccatca ccctgccctg ccgcatcaag cagatcatca acgcccccaa ggccatgtac 1260
```

```
qccccccca tccqcggcca gatccgctgc agcagcaaca tcaccggcct gctgctgacc 1320
cgcgacggcg gcaaggagat cagcaacacc accgagatct tccgccccgg cggcggcgac 1380
atgcgcgaca actggcgcag cgagctgtac aagtacaagg tggtgaagat cgagcccctg 1440
ggcgtggccc ccaccaaggc caagcgccgc gtggtgcagc gcgagaagcg cgccgtgacc 1500
ctgggcgcca tgttcctggg cttcctgggc gccgccggca gcaccatggg cgcccgcagc 1560
ctgaccctga ccgtgcaggc ccgccagctg ctgagcggca tcgtgcagca gcagaacaac 1620
ctgctgcgcg ccatcgaggc ccagcagcac ctgctgcagc tgaccgtgtg gggcatcaag 1680
cagctgcagg cccgcgtgct ggccgtggag cgctacctga aggaccagca gctgctgggc 1740
atctggggct gcagcggcaa gctgatctgc accaccgccg tgccctggaa cgccagctgg 1800
agcaacaaga gcctggacca gatctggaac aacatgacct ggatggagtg ggagcgcgag 1860
atcqacaact acaccaacct gatctacacc ctgatcgagg agagccagaa ccagcaggag 1920
aagaacgagc aggagctgct ggagctggac aagtgggcca gcctgtggaa ctggttcgac 1980
atcagcaagt ggctgtggta catcaagatc ttcatcatga tcgtgggcgg cctggtgggc 2040
ctgcqcatcg tgttcaccgt gctgagcatc gtgaaccgcg tgcgccaggg ctacagcccc 2100
ctgagettee agaceegett eccegeeece egeggeeeeg acegeeeega gggeategag 2160
gaggaggeg gegagegega cegegacege ageageeece tggtgeaegg cetgetggee 2220
ctgatctggg acgacctgcg cagcctgtgc ctgttcagct accaccgcct gcgcgacctg 2280
atcctgateg cegecegeat egtggagetg etgggeegee geggetggga ggeeetgaag 2340
tactggggca acctgctgca gtactggatc caggagctga agaacagcgc cgtgagcctg 2400
ttcgacgcca tcgccatcgc cgtggccgag ggcaccgacc gcatcatcga ggtggcccag 2460
egeateggee gegeetteet geacateece egeegeatee geeagggett egagegegee 2520
ctgctgtaac tcgag
<210> 14
<211> 2529
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Ile424-Ala433
<400> 14
gaattegeea eeatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60
gcagtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccgtgtgga aggaggccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgeacg aggacateat cageetgtgg gaccagagee tgaageeetg egtgaagetg 360
accecectgt gegtgaccet geactgeace aacctgaaga aegecaccaa caccaagage 420
agcaactgga aggagatgga ccgcggcgag atcaagaact gcagcttcaa ggtgaccacc 480
agcatcegea acaagatgea gaaggagtae geeetgttet acaagetgga egtggtgeee 540
atcgacaacg acaacaccag ctacaagctg atcaactgca acaccagcgt gatcacccag 600
gcctgcccca aggtgagctt cgagcccatc cccatccact actgcgcccc cgccggcttc 660
gccatcctga agtgcaacga caagaagttc aacggcagcg gcccctgcac caacgtgagc 720
acceptgeagt geacceaegg cateegeece gtggtgagea eccagetget getgaaegge 780
agectggccg aggagggcgt ggtgatccgc agcgagaact tcaccgacaa cgccaagacc 840
atcatcgtgc agctgaagga gagcgtggag atcaactgca cccgccccaa caacaacacc 900
cgcaagagca tcaccatcgg ccccggccgc gccttctacg ccaccggcga catcatcggc 960
gacatccgcc aggcccactg caacatcagc ggcgagaagt ggaacaacac cctgaagcag 1020
atogtgacca agotgcaggo coagttoggo aacaagacca togtgttoaa goagagcago 1080
ggcggcgacc ccgagategt gatgcacage ttcaactgcg gcggcgagtt cttctactgc 1140
aacagcacco agotgttoaa cagcacotgg aacaacacca toggooccaa caacaccaac 1200
ggcaccatca ccctgccctg ccgcatcaag cagatcatcg gcggcgccat gtacgccccc 1260
cccatcogcg gccagatccg ctgcagcagc aacatcaccg gcctgctgct gacccgcgac 1320
ggcggcaagg agatcagcaa caccaccgag atcttccgcc ccggcggcgg cgacatgcgc 1380
gacaactggc gcagcgagct gtacaagtac aaggtggtga agatcgagcc cctgggcgtg 1440
gececeacea aggeeaageg eegegtggtg eagegegaga agegegeegt gaccetggge 1500
gccatgttcc tgggcttcct gggcgccgcc ggcagcacca tgggcgcccg cagcctgacc 1560
ctgaccgtgc aggcccgcca gctgctgagc ggcatcgtgc agcagcagaa caacctgctg 1620
```

```
eqequateq aggectagea geacetgetg cagetgaceg tgtggggcat caageagetg 1680
caggocogcg tgctggccgt ggagogctac ctgaaggacc agcagotgct gggcatotgg 1740
ggctgcagcg gcaagctgat ctgcaccacc gccgtgccct ggaacgccag ctggagcaac 1800
aagageetgg accagatetg gaacaacatg acctggatgg agtgggageg egagategae 1860
aactacacca acctgatcta caccctgatc gaggagagcc agaaccagca ggagaagaac 1920
gagcaggagc tgctggagct ggacaagtgg gccagcctgt ggaactggtt cgacatcagc 1980
aagtggctgt ggtacatcaa gatcttcatc atgatcgtgg gcggcctggt gggcctgcgc 2040
atogtgttca cogtgctgag catogtgaac cgcgtgcgcc agggctacag ccccctgagc 2100
ttccagacce getteecege eccegegge eccgacegee ecgagggeat egaggaggag 2160
qqqqqqqqq qcqaccqcqa ccqcaqcaqc cccctggtgc acggcctgct ggccctgatc 2220
tgggacgacc tgcgcagcct gtgcctgttc agctaccacc gcctgcgcga cctgatcctg 2280
atogoogooc gcatogtgga gotgotgggo ogoogoggot gggaggooct gaagtactgg 2340
ggcaacctgc tgcagtactg gatccaggag ctgaagaaca gcgccgtgag cctgttcgac 2400
gccatcgcca tcgccgtggc cgagggcacc gaccgcatca tcgaggtggc ccagcgcatc 2460
ggccgcgct tectgcacat eccegecge atecgccagg gettegageg egecetgetg 2520
taactcgag
<210> 15
<211> 2523
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Ile423-Met434
<400> 15
gaattegeea eeatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60
gcagtetteg tttegeceag egeegtggag aagetgtggg tgaeegtgta etaeggegtg 120
cccgtgtgga aggaggccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagategtge tggagaaegt gacegagaae tteaacatgt ggaagaaeaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360
accccctgt gcgtgaccct gcactgcacc aacctgaaga acgccaccaa caccaagagc 420
agcaactgga aggagatgga ccgcggcgag atcaagaact gcagcttcaa ggtgaccacc 480
agcatccgca acaagatgca gaaggagtac gccctgttct acaagctgga cgtggtgccc 540
ategacaacg acaacaccag ctacaagetg ateaactgca acaccagegt gateacccag 600
gcctgcccca aggtgagctt cgagcccatc cccatccact actgcgcccc cgccggcttc 660
gccatcctga agtgcaacga caagaagttc aacggcagcg gcccctgcac caacgtgagc 720
acceptgcagt gcacccacgg catccgcccc gtggtgagca cccagctgct gctgaacggc 780
agectggeeg aggagggegt ggtgateege agegagaaet teacegaeaa egecaagaee 840
atcatcgtgc agetgaagga gagegtggag atcaactgca eeegeeecaa caacaacace 900
cqcaaqaqca tcaccatcgg ccccggccgc gccttctacg ccaccggcga catcatcggc 960
gacatccgcc aggcccactg caacatcagc ggcgagaagt ggaacaacac cctgaagcag 1020
atogtgacca agotgcaggo coagttoggo aacaagacca togtgttcaa goagagcago 1080
ggcggcgacc ccgagatcgt gatgcacagc ttcaactgcg gcggcgagtt cttctactgc 1140
aacagcaccc agetgttcaa cagcacctgg aacaacacca tcggccccaa caacaccaac 1200
qqcaccatca ccctqccctq ccgcatcaag cagatcggcg gcatgtacgc cccccccatc 1260
cgcggccaga tccgctgcag cagcaacatc accggcctgc tgctgacccg cgacggcggc 1320
aaggagatca gcaacaccac cgagatcttc cgccccggcg gcggcgacat gcgcgacaac 1380
tggcgcagcg agctgtacaa gtacaaggtg gtgaagatcg agcccctggg cgtggccccc 1440
accaaggcca agegcegegt ggtgcagege gagaagegeg cegtgaceet gggegecatg 1500
tteetggget teetgggege egeeggeage accatgggeg eeegeageet gaeeetgaee 1560
gtgcaggccc gccagctgct gagcggcatc gtgcagcagc agaacaacct gctgcgcgcc 1620
atcgaggccc agcagcacct gctgcagctg accgtgtggg gcatcaagca gctgcaggcc 1680
cgcgtgctgg ccgtggagcg ctacctgaag gaccagcagc tgctgggcat ctggggctgc 1740
ageggeaage tgatetgeae cacegeegtg ceetggaaeg ceagetggag caacaagage 1800
ctggaccaga tctggaacaa catgacctgg atggagtggg agcgcgagat cgacaactac 1860
accaacctga totacacct gatcgaggag agccagaacc agcaggagaa gaacgagcag 1920
gagetgetgg agetggaeaa gtgggeeage etgtggaaet ggttegaeat cageaagtgg 1980
```

```
ctgtggtaca tcaagatctt catcatgatc gtgggcggcc tggtgggcct gcgcatcgtg 2040
ttcaccgtgc tgagcatcgt gaaccgcgtg cgccagggct acagccccct gagcttccag 2100
accegettee eegeeeeeg eggeeeegae egeeeegagg geategagga ggagggegge 2160
gagegegace gegacegeag cagececetg gtgcaeggee tgctggeeet gatetgggae 2220
gacctgcgca gcctgtgcct gttcagctac caccgcctgc gcgacctgat cctgatcgcc 2280
gcccgcatcg tggagctgct gggccgccgc ggctgggagg ccctgaagta ctggggcaac 2340
ctgctgcagt actggatcca ggagctgaag aacagcgccg tgagcctgtt cgacgccatc 2400
gccatcgccg tggccgaggg caccgaccgc atcatcgagg tggcccagcg catcggccgc 2460
geetteetge acateceeeg eegeateege cagggetteg agegegeeet getgtaacte 2520
gag
<210> 16
<211> 2517
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Gln422-Tyr435
<400> 16
gaattegeea ceatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60
gcagtctteg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccgtgtgga aggaggccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180
accqaqqtqc acaacqtqtq ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagategtge tggagaaegt gaeegagaae tteaacatgt ggaagaaeaa catggtggag 300
cagatgeacg aggacateat cageetgtgg gaccagagee tgaageeetg egtgaagetg 360
accccctgt gcgtgaccct gcactgcacc aacctgaaga acgccaccaa caccaagagc 420
agcaactgga aggagatgga ccgcggcgag atcaagaact gcagcttcaa ggtgaccacc 480
agcateegea acaagatgea gaaggagtae geeetgttet acaagetgga egtggtgeee 540
ategacaacg acaacaccag ctacaagetg atcaactgca acaccagegt gatcacccag 600
geotgeocca aggtgagett egageocate eccatecaet aetgegeece egeoggette 660
gccatcctga agtgcaacga caagaagttc aacggcagcg gcccctgcac caacgtgagc 720
acceptgcagt gcacccacgg catccgccc gtggtgagca cccagctgct gctgaacggc 780
agcottggccg aggaggcgt ggtgatccgc agcgagaact tcaccgacaa cgccaagacc 840
atcatcgtgc agctgaagga gagcgtggag atcaactgca cccgccccaa caacaacacc 900
cqcaaqaqca tcaccatcgg ccccggccgc gccttctacg ccaccggcga catcatcggc 960
gacatccgcc aggcccactg caacatcagc ggcgagaagt ggaacaacac cctgaagcag 1020
atcgtgacca agctgcaggc ccagttcggc aacaagacca tcgtgttcaa gcagagcagc 1080
ggcggcgacc ccgagatcgt gatgcacagc ttcaactgcg gcggcgagtt cttctactgc 1140
aacagcaccc agctgttcaa cagcacctgg aacaacacca tcggccccaa caacaccaac 1200
ggcaccatca ccctgccctg ccgcatcaag cagggcggct acgcccccc catccgcggc 1260
cagateeget geageageaa cateaeegge etgetgetga eeegegaegg eggeaaggag 1320
atcagcaaca ccaccgagat cttccgcccc ggcggcggcg acatgcgcga caactggcgc 1380
agcgagctgt acaagtacaa ggtggtgaag atcgagcccc tggggggtggc ccccaccaag 1440
gccaagegee gegtggtgca gegegagaag egegeegtga eeetgggege catgtteetg 1500
ggetteetgg gegeegeegg cageaceatg ggegeeegea geetgaeeet gaeegtgeag 1560
geoegecage tgetgagegg categtgeag cageagaaca acetgetgeg egecategag 1620
gcccagcagc acctgctgca gctgaccgtg tggggcatca agcagctgca ggcccgcgtg 1680
ctggccgtgg agcgctacct gaaggaccag cagctgctgg gcatctgggg ctgcagcggc 1740
aagctgatct gcaccaccgc cgtgccctgg aacgccagct ggagcaacaa gagcctggac 1800
cagatetgga acaacatgac etggatggag tgggagegeg agategacaa etacaceaac 1860
ctgatctaca ccctgatcga ggagagccag aaccagcagg agaagaacga gcaggagctg 1920
ctggagctgg acaagtgggc cagcctgtgg aactggttcg acatcagcaa gtggctgtgg 1980
tacatcaaga tetteateat gategtggge ggeetggtgg geetgegeat egtgtteace 2040
gtgctgagca tcgtgaaccg cgtgcgccag ggctacagcc ccctgagctt ccagacccgc 2100
ttccccgccc cccgcggccc cgaccgcccc gagggcatcg aggaggaggg cggcgagcgc 2160
gaccgcgacc gcagcagccc cctggtgcac ggcctgctgg ccctgatctg ggacgacctg 2220
cgcagcctgt gcctgttcag ctaccaccgc ctgcgcgacc tqatcctqat cgccqcccqc 2280
ategtggage tgetgggeeg eegeggetgg gaggeeetga agtaetgggg caacetgetg 2340
```

```
cagtactgga tccaggagct gaagaacagc gccgtgagcc tgttcgacgc catcgccatc 2400
geogtggeeg agggeacega cegeateate gaggtggeee agegeategg cegegeette 2460
ctgcacatcc cccgccgcat ccgccagggc ttcgagcgcg ccctgctgta actcgag
<210> 17
<211> 2517
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Gln422-Tyr435B
gaattegeca ccatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60
geagtetteg tttegeecag egeegtggag aagetgtggg tgaeegtgta etaeggegtg 120
cccgtgtgga aggaggccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtge acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360
accecettt gegtgaeeet geaetgeaee aacetgaaga aegeeaeeaa eaceaagage 420
aqcaactqqa aqqaqatqqa ccqcqqcqaq atcaaqaact qcaqcttcaa qqtqaccacc 480
agcatccgca acaagatgca gaaggagtac gccctgttct acaagctgga cgtggtgccc 540
ategacaacg acaacaccag ctacaagctg atcaactgca acaccagcgt gateacccag 600
geetgeecea aggtgagett egageecate cecatecaet aetgegeece egeeggette 660
gccatcctga agtgcaacga caagaagttc aacggcagcg gcccctgcac caacgtgagc 720
acceptqcaqt qcacccacqg catccqcccc gtggtqaqca cccaqctqct qctqaacqqc 780
agcctggccg aggagggcgt ggtgatccgc agcgagaact tcaccgacaa cgccaagacc 840
atcategtgc agetgaagga gagegtggag atcaactgca ceegeeccaa caacaacace 900
cgcaagagca tcaccatcgg ccccggccgc gccttctacg ccaccggcga catcatcggc 960
gacatccgcc aggcccactg caacatcagc ggcgagaagt ggaacaacac cctgaagcag 1020
atcgtgacca agctgcaggc ccagttcggc aacaagacca tcgtgttcaa gcagagcagc 1080
ggcggcgacc ccgagatcgt gatgcacagc ttcaactgcg gcggcgagtt cttctactgc 1140
aacagcaccc agctgttcaa cagcacctgg aacaacacca teggccccaa caacaccaac 1200
ggcaccatca coctgocctg cogcatcaag caggocccct acgcccccc catcogcggc 1260
cagateeget geageageaa cateaeegge etgetgetga eeegegaegg eggeaaggag 1320
atcagcaaca ccaccgagat cttccgccc ggcggcggcg acatgcgcga caactggcgc 1380
agegagetgt acaagtacaa ggtggtgaag ategageeee tgggegtgge ceceaceaag 1440
gccaagegcc gcgtggtgca gcgcgagaag cgcgccgtga ccctgggcgc catgttcctg 1500
ggetteetgg gegeegeegg cageaceatg ggegeeegea geetgaeeet gaeegtgeag 1560
gcccgccagc tgctgagcgg catcgtgcag cagcagaaca acctgctgcg cgccatcgag 1620
geccageage acetgetgea getgacegtg tggggcatea ageagetgea ggecegegtg 1680
ctggccgtgg agcgctacct gaaggaccag cagctgctgg gcatctgggg ctgcagcggc 1740
aagctgatct gcaccaccgc cgtgccctgg aacgccagct ggagcaacaa gagcctggac 1800
cagatotgga acaacatgac otggatggag tgggagcgcg agatogacaa otacaccaac 1860
ctgatctaca ccctgatcga ggagagccag aaccagcagg agaagaacga gcaggagctg 1920
ctggagctgg acaagtgggc cagcctgtgg aactggttcg acatcagcaa gtggctgtgg 1980
tacatcaaga tottoatcat gatogtgggc ggcotggtgg gcotgcgcat cgtgttcacc 2040
gtgctgagca tcgtgaaccg cgtgcgccag ggctacagcc ccctgagctt ccagacccgc 2100
ttccccgccc cccgcggccc cgaccgcccc gagggcatcg aggaggaggg cggcgagcgc 2160
gaccgcgacc gcagcagccc cctggtgcac ggcctgctgg ccctgatctg ggacgacctg 2220
cgcagcetgt gcctgttcag ctaccaccgc ctgcgcgacc tgatcctgat cgccgcccgc 2280
atcgtggagc tgctgggccg ccgcggctgg gaggccctga agtactgggg caacctgctg 2340
cagtactgga tccaggagct gaagaacagc gccgtgagcc tqttcqacqc catcqccatc 2400
gccgtggccg agggcaccga ccgcatcatc gaggtggccc agcgcatcgg ccgcgccttc 2460
ctgcacatcc cccgccgcat ccgccagggc ttcgagcgcg ccctgctgta actcgag
<210> 18
<211> 2322
<212> DNA
```

```
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Leu122-Ser199;
      Arg426-Gly431
gaattegeca ceatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60
gcagtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccgtgtgga aggaggccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagectgtgg gaccagagee tgaageeetg egtgaagetg 360
qqcaacagcg tgatcaccca ggcctgcccc aaggtgagct tcgagcccat ccccatccac 420
tactgcgccc ccgccggctt cgccatcctg aagtgcaacg acaagaagtt caacggcagc 480
ggcccctgca ccaacgtgag caccgtgcag tgcacccacg gcatccgccc cgtggtgagc 540
acccagetge tgetgaacgg cageetggee gaggagggeg tggtgateeg cagegagaac 600
ttcaccqaca acqccaagac catcatcgtg cagctgaagg agagcgtgga gatcaactgc 660
accogceca acaacaacac ecqcaagage atcaccateg geoeggeeg egeettetac 720
gccaccggcg acatcatcgg cgacatccgc caggcccact gcaacatcag cggcgagaag 780
tggaacaaca ccctgaagca gatcgtgacc aagctgcagg cccagttcgg caacaagacc 840
atogtgttca agcagagcag cggcggcgac cccgagatcg tgatgcacag cttcaactgc 900
ggcggcgagt tottotactg caacagcacc cagotgttca acagcacctg gaacaacacc 960
ateggececa acaacaccaa eggeaccate accetgecet geegeateaa geagateate 1020
aaccgeggeg geggeaagge catgtacgee ecceecatee geggeeagat eegetgeage 1080
agcaacatca ccggcctgct gctgacccgc gacggcggca aggagatcag caacaccacc 1140
gagatettee geeeggegg eggegacatg egegacaact ggegeagega getgtacaag 1200
tacaaggtgg tgaagatcga gcccctgggc gtggcccca ccaaggccaa gcgccgcgtg 1260
gtgcagcgcg agaagcgcgc cgtgaccctg ggcgccatgt tcctgggctt cctgggcgcc 1320
geeggeagea ceatgggege cegeageetg accetgaceg tgeaggeeeg ceagetgetg 1380
aggggcateg tgcagcagca gaacaacctg ctgcgcgcca tcgaggccca gcagcacctg 1440
ctgcagctga ccgtgtgggg catcaagcag ctgcaggccc gcgtgctggc cgtggagcgc 1500
tacctgaagg accagcagct gctgggcatc tggggctgca gcggcaagct gatctgcacc 1560
accgccgtgc cctggaacgc cagctggagc aacaagagcc tggaccagat ctggaacaac 1620
atgacctgga tggagtggga gcgcgagatc gacaactaca ccaacctgat ctacaccctg 1680
atcgaggaga gccagaacca gcaggagaag aacgagcagg agctgctgga gctggacaag 1740
tgggccagcc tgtggaactg gttcgacatc agcaagtggc tgtggtacat caagatctic 1800
atcatgateg tgggcggcct ggtgggcctg cgcatcgtgt tcaccgtgct gagcatcgtg 1860
aaccgcgtgc gccagggcta cagcccctg agcttccaga cccgcttccc cgcccccgc 1920
ggccccgacc gccccgaggg catcgaggag gagggcggcg agcgcgaccg cgaccgcagc 1980
agccccctqq tqcacqqcct gctgqccctq atctggqacq acctqcqcaq cctqtqcctq 2040
ttcagctacc accgcctgcg cgacctgatc ctgatcgccg cccgcatcgt ggagctgctg 2100
ggccgccgcg gctgggaggc cctgaagtac tggggcaacc tgctgcagta ctggatccag 2160
gagetgaaga acagegeegt gageetgtte gaegecateg ceategeegt ggeegaggge 2220
accgaccgca tcatcgaggt ggcccagcgc atcggccgcq ccttcctgca catcccccqc 2280
cgcatccgcc agggcttcga gcgcgccctg ctgtaactcg ag
<210> 19
<211> 2322
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Leu122-Ser199;
     Arg426-Lys432
<400> 19
gaattegeea ceatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60
gcagtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
```

```
cccgtgtgga aggaggccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagatogtgo tggagaacgt gacogagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360
ggcaacagcg tgatcaccca ggcctgcccc aaggtgagct tcgagcccat ccccatccac 420
tactqcqccc ccqccggctt cgccatcctg aagtgcaacg acaagaagtt caacggcagc 480
ggcccctgca ccaacgtgag caccgtgcag tgcacccacg gcatccgccc cgtggtgagc 540
accoagetge tgetgaacgg cageetggee gaggagggeg tggtgateeg cagegagaac 600
ttcaccqaca acgccaagac catcatcgtg cagctgaagg agagcgtgga gatcaactgc 660
accogcocca acaacaacac cogcaagago atcaccatog goocoggoog ogcottotac 720
gccaccggcg acatcatcgg cgacatccgc caggcccact gcaacatcag cggcgagaag 780
tggaacaaca coctgaagca gatcgtgacc aagctgcagg cccagttcgg caacaagacc 840
atcgtgttca agcagagcag cggcggcgac cccgagatcg tgatgcacag cttcaactgc 900
ggcggcgagt tcttctactg caacagcacc cagctgttca acagcacctg gaacaacacc 960
ateggececa acaacaceaa eggeaceate accetgecet geegeateaa geagateate 1020
aaccgcggcg gcaacaaggc catgtacgcc cccccatcc gcggccagat ccgctgcagc 1080
agcaacatca ccggcctgct gctgacccgc gacggcggca aggagatcag caacaccacc 1140
gagatettee geeeggegg eggegacatg egegacaact ggegeagega getgtacaag 1200
tacaaggtgg tgaagatcga gcccctgggc gtggcccca ccaaggccaa gcgccgcgtg 1260
gtgcagcgcg agaagcgcgc cgtgaccctg ggcgccatgt tcctgggctt cctgggcgcc 1320
geoggeagea ceatgggege eegeageetg accetgaceg tgeaggeeeg ceagetgetg 1380
ageggcateg tgcagcagca gaacaacetg etgcgcgcca tcgaggccca gcagcacetg 1440
ctgcagctga ccgtgtgggg catcaagcag ctgcaggccc gcgtgctggc cgtggagcgc 1500
tacctgaagg accagcagct gctgggcatc tggggctgca gcggcaagct gatctgcacc 1560
accgccgtgc cctggaacgc cagctggagc aacaagagcc tggaccagat ctggaacaac 1620
atgacctgga tggagtggga gcgcgagatc gacaactaca ccaacctgat ctacaccctg 1680
atcgaggaga gccagaacca gcaggagaag aacgagcagg agctgctgga gctggacaag 1740
tgggccagcc tgtggaactg gttcgacatc agcaagtggc tgtggtacat caagatcttc 1800
atcatgateg tgggcgcct ggtgggcctg cgcatcgtgt tcaccgtgct gagcatcgtg 1860
aaccgcgtgc gccagggcta cagcccctg agcttccaga cccgcttccc cgcccccgc 1920
ggccccgacc gccccgaggg catcgaggag gagggcggcg agcgcgaccg cgaccgcagc 1980
agccccctgg tgcacggcct gctggccctg atctgggacg acctgcgcag cctgtgcctg 2040
ttcagctacc accgcctgcg cgacctgatc ctgatcgccg cccgcatcgt ggagctgctg 2100
ggccgccgcg gctgggaggc cctgaagtac tggggcaacc tgctgcagta ctggatccag 2160
gagetgaaga acagegeegt gageetgtte gaegeeateg ceategeegt ggeegaggge 2220
accgaccgca tcatcgaggt ggcccagcgc atcggccgcg ccttcctgca catcccccgc 2280
cgcatccgcc agggcttcga gcgcgccctg ctgtaactcg ag
<210> 20
<211> 2322
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Leu122-Ser199;
      Trp427-Gly431
<400> 20
gaattegeea eeatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60
gcagtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcqtg 120
cccgtgtgga aggaggccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctq 360
ggcaacagcg tgatcaccca ggcctgcccc aaggtgagct tcgagcccat ccccatccac 420
tactgcgccc ccgccggctt cgccatcctg aagtgcaacg acaagaagtt caacggcagc 480
ggcccctgca ccaacgtgag caccgtgcag tgcacccacg gcatccgccc cgtggtgagc 540
acceagetge tgctgaacgg cageetggee gaggagggeg tggtgateeg cagegagaac 600
ttcaccgaca acgccaagac catcatcgtg cagctgaagg agagcgtgga gatcaactgc 660
```

```
accegeecca acaacaacac cegeaagage ateaceateg geeceggeeg egeettetac 720
gccaccggcg acatcatcgg cgacatccgc caggcccact gcaacatcag cggcgagaag 780
tggaacaaca coctgaagca gatcgtgacc aagctgcagg cccagttcgg caacaagacc 840
atogtgttca agcagagcag cggcggcgac cccgagatcg tgatgcacag cttcaactgc 900
ggcggcgagt tottotactg caacagcacc cagotgttca acagcacctg gaacaacacc 960
atoggococa acaacaccaa oggoaccato accotgocot geogoatcaa geagateato 1020
aaccgctggg gcggcaaggc catgtacgcc cccccatcc gcggccagat ccgctgcagc 1080
agcaacatca ceggeetget getgaeeege gaeggeggea aggagateag caacaccace 1140
gagatettee geceeggegg eggegacatg egegacaaet ggegeagega getgtacaag 1200
tacaaggtgg tgaagatcga gcccctgggc gtggccccca ccaaggccaa gcgccgcgtg 1260
gtgcagcgcg agaagcgcgc cgtgaccctg ggcgccatgt tcctgggctt cctgggcgcc 1320
geoggeagea ceatgggege cegeageetg accetgaceg tgeaggeeeg ceagetgetg 1380
ageggeateg tgeageagea gaacaacetg etgegegeea tegaggeeea geageacetg 1440
ctgcagctga ccgtgtgggg catcaagcag ctgcaggccc gcgtgctggc cgtggagcgc 1500
tacctgaagg accagcagct gctgggcatc tggggctgca gcggcaagct gatctgcacc 1560
accgccgtgc cctggaacgc cagctggagc aacaagagcc tggaccagat ctggaacaac 1620
atgacctgga tggagtggga gcgcgagatc gacaactaca ccaacctgat ctacaccetg 1680
atcgaggaga gccagaacca gcaggagaag aacgagcagg agctgctgga gctggacaag 1740
tgggccagcc tgtggaactg gttcgacatc agcaagtggc tgtggtacat caagatcttc 1800
atcatgatcg tgggcggcct ggtgggcctg cgcatcgtgt tcaccgtgct gagcatcgtg 1860
aaccgcgtgc gccagggcta cagcccctg agcttccaga cccgcttccc cgccccccgc 1920
ggccccgacc gccccgaggg catcgaggag gagggcggcg agcgcgaccg cgaccgcagc 1980
agececetgg tgeaeggeet getggeeetg atetgggaeg acetgegeag cetgtgeetg 2040
ttcagctacc accgcctgcg cgacctgatc ctgatcgccg cccgcatcgt ggagctgctg 2100
ggcegcegcg gctgggaggc cctgaagtac tggggcaacc tgctgcagta ctggatccag 2160
gagetgaaga acagegeegt gageetgtte gaegeeateg ceategeegt ggeegaggge 2220
accqaccgca tcatcgaggt ggcccagcgc atcggccgcg ccttcctgca catcccccgc 2280
cgcatccgcc agggcttcga gcgcgccctg ctgtaactcg ag
                                                                  2322
<210> 21
<211> 2310
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Lys121-Val200;
      Asn425-Lys432
<400> 21
gaattegeea ceatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60
gcagtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccgtgtgga aggaggccac caccaccctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaaggcc 360
cccgtgatca cccaggcctg ccccaaggtg agcttcgagc ccatccccat ccactactgc 420
gccccgccg gcttcgccat cctgaagtgc aacgacaaga agttcaacgg cagcggcccc 480
tgcaccaacg tgagcaccgt gcagtgcacc cacggcatcc gccccgtggt gagcacccag 540
ctgctgctga acggcagcct ggccgaggag ggcgtggtga tccgcagcga gaacttcacc 600
gacaacgcca agaccatcat cgtgcagctg aaggagagcg tggagatcaa ctgcacccgc 660
cccaacaaca acaccegcaa gagcatcacc ateggeeceg geegegeett ctaegeeacc 720
ggcgacatca tcggcgacat ccgccaggcc cactgcaaca tcagcggcga gaagtggaac 780
aacaccctga agcagatcgt gaccaagctg caggcccagt tcqqcaacaa qaccatcqtq 840
ttcaagcaga gcagcggcgg cgaccccgag atcgtgatgc acagcttcaa ctgcggcggc 900
gagttettet actgeaacag cacceagetg tteaacagea cetggaacaa caccategge 960
cccaacaaca ccaacggcac catcaccctg ccctgccgca tcaagcagat catcaacgcc 1020
cccaaggcca tgtacgcccc ccccatccgc ggccagatcc gctgcagcag caacatcacc 1080
ggcctgctgc tgacccgcga cggcggcaag gagatcagca acaccaccga gatcttccgc 1140
cccggcggcg gcgacatgcg cgacaactgg cgcagcgagc tgtacaagta caaggtggtg 1200
```

```
aagatcgagc ccctgggcgt ggccccacc aaggccaagc gccgcgtggt gcagcgcgag 1260
aagcgcgccg tgaccctggg cgccatgttc ctgggcttcc tgggcgccgc cggcagcacc 1320
atgggggcc qcagcctgac cctgaccgtg caggcccgcc agctgctgag cggcatcgtg 1380
cagcagcaga acaacctgct gcgcgccatc gaggcccagc agcacctgct gcagctgacc 1440
gtgtggggca tcaagcagct gcaggcccgc gtgctggccg tggagcgcta cctgaaggac 1500
cagcagetge tgggcatetg gggetgeage ggcaagetga tetgeaceae egeegtgeee 1560
tggaacgcca gctggagcaa caagagcctg gaccagatct ggaacaacat gacctggatg 1620
gagtgggagc gcgagatcga caactacacc aacctgatct acaccctgat cgaggagagc 1680
cagaaccagc aggagaagaa cgagcaggag ctgctggagc tggacaagtg ggccagcctg 1740
tggaactggt tcgacatcag caagtggctg tggtacatca agatcttcat catgatcgtg 1800
ggcggcctgg tgggcctgcg catcgtgttc accgtgctga gcatcgtgaa ccgcgtgcgc 1860
cagggetaca geoccetgag ettecagace egetteeceg ecceegegg eccegacege 1920
cccgagggca tcgaggagga gggcggcgag cgcgaccgcg accgcagcag ccccctggtg 1980
caeggeetge tggeeetgat etgggaegae etgegeagee tgtgeetgtt eagetaceae 2040
cgcctgcgcg acctgatect gategccgcc cgcategtgg agetgctggg ccgccgcggc 2100
tgggaggccc tgaagtactg gggcaacctg ctgcagtact ggatccagga gctgaagaac 2160
agegeegtga geetgttega egecategee ategeegtgg eegagggeae egacegeate 2220
ategaggtgg cccagegeat eggeegegee tteetgeaca tecceegeeg cateegeeag 2280
ggcttcgagc gcgccctgct gtaactcgag
<210> 22
<211> 2298
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Val120-Ile201;
      Ile424-Ala433
<400> 22
gaattegeea eeatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60
gcagtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccgtgtgga aggaggccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgggcggc 360
atcacccagg cctgccccaa ggtgagcttc gagcccatcc ccatccacta ctgcgccccc 420
gccggcttcg ccatcctgaa gtgcaacgac aagaagttca acggcagcgg cccctgcacc 480
aacgtgagca ccgtgcagtg cacccacggc atccgccccg tggtgagcac ccagctgctg 540
ctgaacggca gcctggccga ggagggcgtg gtgatccgca gcgagaactt caccgacaac 600
gccaagacca tcatcgtgca gctgaaggag agcgtggaga tcaactgcac ccgccccaac 660
aacaacaccc gcaagagcat caccategge eceggeegeg cettetaege caceggegae 720
atcatcggcg acatccgcca ggcccactgc aacatcagcg gcgagaagtg gaacaacacc 780
ctgaagcaga tcgtgaccaa gctgcaggcc cagttcggca acaagaccat cgtgttcaag 840
cagagcagcg gcggcgaccc cgagatcgtg atgcacagct tcaactgcgg cggcgagttc 900
ttctactgca acagcaccca gctgttcaac agcacctgga acaacaccat cggccccaac 960
aacaccaacg gcaccatcac cctgccctgc cgcatcaagc agatcatcgg cggcgccatg 1020
tacgccccc ccatccgcgg ccagatccgc tgcagcagca acatcaccgg cctgctgctg 1080
accegegacg geggeaagga gateageaac accacegaga tetteegeec eggeggegge 1140
gacatgegeg acaactggeg cagegagetg tacaagtaca aggtggtgaa gategageee 1200
ctgggcgtgg cccccaccaa ggccaagcgc cgcgtggtgc agcgcgagaa gcgcgcgtg 1260
accetgggeg ceatgtteet gggetteetg ggegeegeeg geageaceat gggegeege 1320
agcetgacee tgacegtgea ggeeegeeag etgetgageg geategtgea geageagaae 1380
aacctgctgc gcgccatcga ggcccagcag cacctgctgc agctgaccgt gtggggcatc 1440
aagcagctgc aggcccgcgt gctggccgtg gagcgctacc tgaaggacca gcagctgctg 1500
ggcatctggg gctgcagcgg caagctgatc tgcaccaccg ccgtgccctg gaacgccagc 1560
tggagcaaca agagcctgga ccagatctgg aacaacatga cctggatgga gtgggagcgc 1620
gagategaca actacaceaa cetgatetae accetgateg aggagageea gaaceageag 1680
gagaagaacg agcaggagct gctggagctg gacaagtggg ccagcctgtg gaactggttc 1740
```

```
gacatcagca agtggctgtg gtacatcaag atcttcatca tgatcgtggg cggcctggtg 1800
ggcctgcgca tcgtgttcac cgtgctgagc atcgtgaacc gcgtgcgcca gggctacagc 1860
cccctgagct tccagacccg cttccccgcc ccccgcggcc ccgaccgccc cgagggcatc 1920
gaggaggagg geggegageg egacegegae egeageagee ecetggtgea eggeetgetg 1980
gccctgatct gggacgacct gcgcagcctg tgcctgttca gctaccaccg cctgcgcgac 2040
ctgatcctga tcgccgcccg catcgtggag ctgctgggcc gccgcggctg ggaggccctg 2100
aagtactggg gcaacctgct gcagtactgg atccaggagc tgaagaacag cgccgtgagc 2160
ctgttcgacg ccatcgccat cgccgtggcc gagggcaccg accgcatcat cgaggtggcc 2220
cagegeateg geogegeett cetgeacate eccegeegea teegeeaggg ettegagege 2280
gccctgctgt aactcgag
<210> 23
<211> 2298
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence:
      Val120-Ile201B; Ile424-Ala433
<400> 23
gaattegeea ceatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60
qcaqtcttcq tttcqcccaq cqccqtqqaq aaqctqtqqq tqaccqtqta ctacqqcqtq 120
cccgtgtgga aggaggccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180
acegaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagatogtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgcccggc 360
atcacccagg cctgccccaa ggtgagcttc gagcccatcc ccatccacta ctgcgccccc 420
geoggetteg ceatectgaa gtgcaacgae aagaagttea aeggeagegg ceeetgeace 480
aacgtgagca cogtgcagtg cacccacggc atccgccccg tggtgagcac ccagctgctg 540
ctgaacggca gcctggccga ggagggcgtg gtgatccgca gcgagaactt caccgacaac 600
gccaagacca tcatcgtgca gctgaaggag agcgtggaga tcaactgcac ccgccccaac 660
aacaacaccc gcaagagcat caccatcggc cccggccgcg ccttctacgc caccggcgac 720
atcateggeg acateegeca ggeecactge aacateageg gegagaagtg gaacaacace 780
ctgaagcaga tcgtgaccaa gctgcaggcc cagttcggca acaagaccat cgtgttcaag 840
cagageageg geggegaccc egagategtg atgeacaget teaactgegg eggegagtte 900
ttctactgca acagcaccca gctgttcaac agcacctgga acaacaccat cggccccaac 960
aacaccaacg gcaccatcac cctgccctgc cgcatcaagc agatcatcgg cggcgccatg 1020
tacgccccc ccatccgcgg ccagatccgc tgcagcagca acatcaccgg cctgctgctg 1080
accegegacg geggcaagga gateagcaac accacegaga tetteegeec eggeggegge 1140
gacatgegeg acaactggeg cagegagetg tacaagtaca aggtggtgaa gategagece 1200
ctgggcgtgg cccccaccaa ggccaagcgc cgcgtggtgc agcgcgagaa gcgcgccgtg 1260
accetgggeg coatgtteet gggetteetg ggegeegeeg geageaceat gggegeeege 1320
agcotgacco tgacogtgca ggcccgccag ctgctgagcg gcatcgtgca gcagcagaac 1380
aacctgctgc gcgccatcga ggcccagcag cacctgctgc agctgaccgt gtggggcatc 1440
aagcagetge aggeeegegt getggeegtg gagegetace tgaaggacea geagetgetg 1500
ggcatctggg gctgcagcgg caagctgatc tgcaccaccg ccgtgccctg gaacgccagc 1560
tggagcaaca agagcctgga ccagatctgg aacaacatga cctggatgga gtgggagcgc 1620
gagategaca actacaceaa cetgatetae accetgateg aggagageea gaaceageag 1680
gagaagaacg agcaggagct gctggagctg gacaagtggg ccagcctgtg gaactggttc 1740
gacatcagca agtggctgtg gtacatcaag atcttcatca tgatcgtggg cggcctggtg 1800
ggcctgcgca tcgtgttcac cgtgctgagc atcgtgaacc gcgtgcgcca gggctacagc 1860
cccctgaget tecagaceeg etteccegee eccegeggee ecgaeegeee egagggeate 1920
gaggaggagg gcggcgagcg cgaccgcgac cgcagcagcc ccctggtgca cggcctgctg 1980
gccctgatct gggacgacct gcgcagcctg tgcctgttca gctaccaccg cctgcgcgac 2040
ctgatectga tegeogeeeg categtggag etgetgggee geegeggetg ggaggeeetg 2100
aagtactggg gcaacctgct gcagtactgg atccaggagc tgaagaacag cgccgtgagc 2160
```

ctgttcgacg ccatcgccat cgccgtggcc gagggcaccg accgcatcat cgaggtggcc 2220 cagcgcatcg gccgcgcctt cctgcacatc ccccgccgca tccgccaggg cttcgagcgc 2280

gccctgctgt aactcgag

2298

```
<210> 24
<211> 2298
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Val120-Thr202;
      Ile424-Ala433
<400> 24
gaattegeea ceatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60
gcagtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccqtqtqqa aggaggccac caccaccctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagategtge tggagaaegt gacegagaae tteaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgggcggc 360
qccacccaqq cctqcccaa qqtqaqcttc qaqcccatcc ccatccacta ctqcqcccc 420
gceggetteg ccateetgaa gtgcaacgac aagaagttea acggcagegg cccetgcace 480
aacgtgagca cegtgcagtg cacccaeggc atcegeeeeg tggtgagcae ecagetgetg 540
ctgaacggca gcctggccga ggagggcgtg gtgatccgca gcgagaactt caccgacaac 600
gccaagacca tcatcgtgca gctgaaggag agcgtggaga tcaactqcac ccqccccaac 660
aacaacacc gcaagagcat caccategge eeeggeegeg eettetaege caceggegae 720
atcateggeg acateegeea ggeecactge aacateageg gegagaagtg gaacaacace 780
ctgaagcaga tcgtgaccaa gctgcaggcc cagttcggca acaagaccat cgtgttcaag 840
cagagcagcg gcggcgaccc cgagatcgtg atgcacagct tcaactgcgg cggcgagttc 900
ttctactgca acagcaccca gctgttcaac agcacctgga acaacaccat cggccccaac 960
aacaccaacg gcaccatcac cctgccctgc cgcatcaagc agatcatcgg cggcgccatg 1020
tacgcccccc ccatccgcgg ccagatccgc tgcagcagca acatcaccgg cctgctgctg 1080
accegegacg geggeaagga gateageaac accacegaga tetteegeec eggeggegge 1140
gacatgcgcg acaactggcg cagcgagctg tacaagtaca aggtggtgaa gatcgagccc 1200
ctgggcgtgg ccccaccaa ggccaagcgc cgcgtggtgc agcgcgagaa gcgcgccgtg 1260
accetgggeg coatgitect gggettectg ggegeegeeg geageaceat gggegeege 1320
agectgacce tgaccgtgca ggcccgccag ctgctgagcg gcatcqtqca qcaqcaqaac 1380
aacctgctgc gcgccatcga ggcccagcag cacctgctgc agctgaccgt gtggggcatc 1440
aagcagctgc aggcccgcgt gctggccgtg gagcgctacc tgaaggacca gcagctgctg 1500
ggcatctggg gctgcagcgg caagctgatc tgcaccaccg ccgtgccctg gaacgccagc 1560
tggagcaaca agagcctgga ccagatctgg aacaacatga cctggatgga gtgggagcgc 1620
gagategaca actacaceaa ectgatetae accetgateg aggagageea gaaceageag 1680
gagaagaacg agcaggagct gctggagctg gacaagtggg ccagcctgtg gaactggttc 1740
gacatcagca agtggctgtg gtacatcaag atcttcatca tgatcgtggg cggcctggtg 1800
ggcctgcgca tcgtgttcac cgtgctgagc atcgtgaacc gcgtgcgcca gggctacagc 1860
cccctgagct tccagacccg cttccccgcc cccgcggcc ccqaccqccc cqaqqqcatc 1920
gaggaggagg gcggcgagcg cgaccgcgac cgcagcagcc ccctggtgca cggcctgctg 1980
geoctgatet gggacgacet gegeageetg tgeetgttea getaceaceg cetgegegae 2040
ctgatcctga tcgccgcccg catcgtggag ctgctgggcc gccgcggctg ggaggccctg 2100
aagtactggg gcaacctgct gcagtactgg atccaggagc tgaaqaacag cgccqtqaqc 2160
ctgttcgacg ccatcgccat cgccgtggcc gagggcaccg accgcatcat cgaggtggcc 2220
cagegeateg geogegeett cetgeacate eccegeegea teegeeaggg ettegagege 2280
gccctgctgt aactcgag
                                                                  2298
<210> 25
<211> 2358
<212> DNA
<213> Artificial Sequence
<223> Description of Artificial Sequence: Vall27-Asn195
```

```
<400> 25
gaattegeca ecatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60
gcagtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccgtgtgga aggaggccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360
acceccetqt qeqtggggc agggaactgc aacaccageg tgatcaccca gqcctqcccc 420
aaggtgaget tegageeeat ecceateeac tactgegeee eegeeggett egeeateetg 480
aagtgcaacg acaagaagtt caacggcagc ggcccctgca ccaacgtgag caccgtgcag 540
tgcacccacg gcatccgccc cgtggtgagc acccagctgc tgctgaacgg cagcctqqcc 600
gaggagggcg tggtgatccg cagcgagaac ttcaccgaca acgccaagac catcatcqtq 660
cagctgaagg agagcgtgga gatcaactgc accegccca acaacaacac ccgcaagagc 720
atcaccateg geoeggeeg egeettetae geoaceggeg acateategg egacateege 780
caggeceact geaacateag eggegagaag tggaacaaca eeetgaagea gategtgaee 840
cccgagatcg tgatgcacag cttcaactgc ggcggcgagt tcttctactg caacagcacc 960
cagctgttca acagcacctg gaacaacacc atcggcccca acaacaccaa cggcaccatc 1020
accetgeeet geegeateaa geagateate aacegetgge aggaggtggg caaggeeatg 1080
tacgccccc ccatccgcgg ccagatccgc tgcagcagca acatcaccgg cctqctqctg 1140
accegegacg geggeaagga gateageaac accacegaga tetteegeec eggeggegge 1200
gacatgcgcg acaactggcg cagcgagctg tacaagtaca aggtggtgaa gatcgagccc 1260
ctgggcgtgg cccccaccaa ggccaagcgc cgcgtggtgc agcgcgagaa gcgcgccgtg 1320
accetgggeg ceatgtteet gggetteetg ggegeegeeg geageaceat gggegeeege 1380
agectgaece tgacegtgea ggeeegeeag etgetgageg geategtgea geageagaae 1440
aacctgctgc gcgccatcga ggcccagcag cacctgctgc agctgaccgt gtggggcatc 1500
aagcagctgc aggcccgcgt gctggccgtg gagcgctacc tgaaggacca gcagctgctg 1560
ggcatctggg gctgcagcgg caagctgatc tgcaccaccg ccgtgccctg gaacgccagc 1620
tggagcaaca agagcctgga ccagatctgg aacaacatga cctggatgga gtgggagcgc 1680
gagategaca actacaceaa cetgatetae accetgateg aggagageea gaaceageag 1740
gagaagaacg agcaggagct gctggagctg gacaagtggg ccagcctgtg gaactggttc 1800
gacatcagca agtggctgtg gtacatcaag atcttcatca tgatcgtggg cggcctggtg 1860
ggcctgcgca tcgtgttcac cgtgctgagc atcgtgaacc gcgtgcgcca gggctacagc 1920
cccctgagct tccagacccg cttccccgcc ccccgcgcc ccgaccgccc cgagggcatc 1980
gaggaggagg gcggcgagcg cgaccgcgac cgcagcagcc ccctggtgca cggcctgctg 2040
gccctgatct gggacgacct gcgcagcctg tgcctgttca gctaccaccg cctgcqcqac 2100
ctgatcctga tegeogeceg categtggag etgetgggee geegeggetg ggaggeeetg 2160
aagtactggg gcaacctgct gcagtactgg atccaggagc tgaagaacag cgccgtgagc 2220
ctgttcgacg ccatcgccat cgccgtggcc gagggcaccg accgcatcat cgaggtggcc 2280
cagegeateg geogegeett cetgeacate eccegeegea teegeeaggg ettegagege 2340
gccctgctgt aactcgag
<210> 26
<211> 2352
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Val127-Asn195;
      Arg426-Gly431
<400> 26
gaattegeea eeatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60
gcagtetteg tttegeccag egeegtggag aagetgtggg tgacegtgta etaeggegtg 120
cccgtgtgga aggaggccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagategtge tggagaaegt gaeegagaae tteaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctq cqtqaaqctq 360
accecetyt gegtggggge agggaactge aacaceageg tgateaceca ggeetgeeec 420
```

	aaggtgagct	tcgagcccat	ccccatccac	tactgcgccc	ccgccggctt	cgccatcctg	480
	aagtgcaacg	acaagaagtt	caacggcagc	ggcccctgca	ccaacgtgag	caccgtgcag	540
	tgcacccacg	gcatccgccc	cgtggtgagc	acccagctgc	tgctgaacgg	cagcctggcc	600
	gaggagggcg	tggtgatccg	cagcgagaac	ttcaccgaca	acgccaagac	catcatcgtg	660
	cagctgaagg	agagcgtgga	gatcaactgc	acccgcccca	acaacaacac	ccgcaagagc	720
	atcaccatcg	gccccggccg	cgccttctac	gccaccggcg	acatcatcgg	cgacatccgc	780
	caggcccact	gcaacatcag	cggcgagaag	tggaacaaca	ccctgaagca	gatcgtgacc	840
	aagctgcagg	cccagttcgg	caacaagacc	atcgtgttca	agcagagcag	cggcggcgac	900
	cccgagatcg	tgatgcacag	cttcaactgc	ggcggcgagt	tcttctactg	caacagcacc	960
	cagctgttca	acagcacctg	gaacaacacc	atcggcccca	acaacaccaa	cggcaccatc	1020
	accctgccct	gccgcatcaa	gcagatcatc	aaccgcggcg	gcggcaaggc	catgtacgcc	1080
	cccccatcc	gcggccagat	ccgctgcagc	agcaacatca	ccggcctgct	gctgacccgc	1140
	gacggcggca	aggagatcag	caacaçcacc	gagatcttcc	gccccggggg	cggcgacatg	1200
	cgcgacaact	ggcgcagcga	gctgtacaag	tacaaggtgg	tgaagatcga	gcccctgggc	1260
	gtggccccca	ccaaggccaa	gcgccgcgtg	gtgcagcgcg	agaagcgcgc	cgtgaccctg	1320
	ggcgccatgt	tcctgggctt	cctgggcgcc	gccggcagca	ccatgggcgc	ccgcagcctg	1380
	accctgaccg	tgcaggcccg	ccagctgctg	agcggcatcg	tgcagcagca	gaacaacctg	1440
	ctgcgcgcca	tcgaggccca	gcagcacctg	ctgcagctga	ccgtgtgggg	catcaagcag	1500
	ctgcaggccc	gcgtgctggc	cgtggagcgc	tacctgaagg	accagcagct	gctgggcatc	1560
	tggggctgca	gcggcaagct	gatctgcacc	accgccgtgc	cctggaacgc	cagctggagc	1620
•	aacaagagcc	tggaccagat	ctggaacaac	atgacctgga	tggagtggga	gcgcgagatc	1680
	gacaactaca	ccaacctgat	ctacaccctg	atcgaggaga	gccagaacca	gcaggagaag	1740
	aacgagcagg	agctgctgga	gctggacaag	tgggccagcc	tgtggaactg	gttcgacatc	1800
	agcaagtggc	tgtggtacat	caagatcttc	atcatgatcg	tgggcggcct	ggtgggcctg	1860
	cgcatcgtgt	tcaccgtgct	gagcatcgtg	aaccgcgtgc	gccagggcta	cagccccctg	1920
	agcttccaga	cccgcttccc	cgcccccgc	ggccccgacc	gccccgaggg	catcgaggag	1980
	gagggcggcg	agcgcgaccg	cgaccgcagc	agccccctgg	tgcacggcct	gctggccctg	2040
	atctgggacg	acctgcgcag	cctgtgcctg	ttcagctacc	accgcctgcg	cgacctgatc	2100
	ctgatcgccg	cccgcatcgt	ggagctgctg	ggccgccgcg	gctgggaggc	cctgaagtac	2160
	tggggcaacc	tgctgcagta	ctggatccag	gagctgaaga	acagcgccgt	gagcctgttc	2220
	gacgccatcg	ccatcgccgt	ggccgagggc	accgaccgca	tcatcgaggt	ggcccagcgc	2280
	atcggccgcg	ccttcctgca	catcccccgc	cgcatccgcc	agggcttcga	gcgcgccctg	2340
	ctgtaactcg	ag					2352